



QY 727 TACACTGCAATCTGTCGGGAGTTTGAGTTTGGAGCGCATCGCCCTTTGTGAGAGACG 786  
Db |||||  
QY 379 TACACTGCAATCTGTCGGGAGTTTGAGTTTGGAGCGCATCGCCCTTTGTGAGAGACG 438  
Db |||||  
QY 787 AGCGGGCTGATCCCTTAAGATC 810  
Db |||||  
QY 439 AGCGGGCTGATCCCTTAAGATC 462

## RESULT 9

AA886762  
ID AA886762 standard; cDNA; 621 BP.

AC AA886762;  
XX

DT 13-FEB-2002 (first entry)  
XX

DE DNA encoding novel human diagnostic protein #22566.  
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX

OS Homo sapiens.  
XX

PN WO200175067-A2.  
XX

PD 11-OCT-2001.  
XX

PF 30-MAR-2001; 2001WO-US08631.  
XX

PR 31-MAR-2000; 2000US-0540217.  
XX

PR 23-AUG-2000; 2000US-0649167.  
XX

PA (HYSE-) HYSEQ INC.  
XX

PI Drmanac RT, Liu C, Tang YT;  
XX

DR WPI; 2001-639362/73.  
XX

DR P-PSDB; ABG22575.  
XX

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
XX

PS Claim 1; SEQ ID No 22566; 103pp; English.  
XX

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA886762 represents novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 621 BP; 156 A; 155 C; 175 G; 135 T; 0 other;  
XX

Query Match

17.3%; Score 218; DB 23; Length 621;

Best Local Similarity 100.0%; Pred. No. 1e-95;  
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 595 ATTACGAGTATCGGATGCGCAGGAGTGGAGAGCCCTTTTCAGGGGCGCTGCGAG 654  
Db |||||

QY 58 ATTACGAGTATCGGATGCGCAGGAGTGGAGAGCCCTTTTCAGGGGCGCTGCGAG 117  
Db |||||

QY 655 TGAATGGCAGCAAGGACCTGCGAGGACGTGTCCATCACTGTCTCAACGTCACCTGAC 714  
Db |||||

QY 118 TGAATGGCAGCAAGGACCTGCGAGGACGTGTCCATCACTGTCTCAACGTCACCTGAC 177  
Db |||||

QY 715 GACTCTGGCCTTACACCTGCAATGTGTCGCGGAGTTTGAAGTTGAGGCGCATCGGCC 774  
Db |||||

QY 178 GACTCTGGCCTTACACCTGCAATGTGTCGCGGAGTTTGAAGTTGAGGCGCATCGGCC 237  
Db |||||

QY 775 TTTGTGAAGACGACGCGGCTGATCCCTTAAGATC 812  
Db |||||

QY 238 TTTGTGAAGACGACGCGGCTGATCCCTTAAGATC 275  
Db |||||

## RESULT 10

AA886762

ID AA886762 standard; DNA; 4625 BP.

AC AA886761;  
XX

DT 21-NOV-2001 (first entry)  
XX

DE Human reproductive system related antigen DNA SEQ ID NO: 7659.  
XX

KW Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.  
XX

OS Homo sapiens.  
XX

PN WO200155320-A2.  
XX

PD 02-AUG-2001.  
XX

PF 17-JAN-2001; 2001WO-US01339.  
XX

PR 31-JAN-2000; 2000US-0179065.  
XX

PR 04-FEB-2000; 2000US-0180628.  
XX

PR 24-FEB-2000; 2000US-0184664.  
XX

PR 02-MAR-2000; 2000US-0186350.  
XX

PR 16-MAR-2000; 2000US-0189874.  
XX

PR 17-MAR-2000; 2000US-0190076.  
XX

PR 18-APR-2000; 2000US-0198123.  
XX

PR 19-MAY-2000; 2000US-0205515.  
XX

PR 07-JUN-2000; 2000US-0209467.  
XX

PR 28-JUN-2000; 2000US-0214886.  
XX

PR 30-JUN-2000; 2000US-0215135.  
XX

PR 07-JUL-2000; 2000US-0216647.  
XX

PR 11-JUL-2000; 2000US-0216880.  
XX

PR 11-JUL-2000; 2000US-0217487.  
XX

PR 14-JUL-2000; 2000US-0217496.  
XX

PR 26-JUL-2000; 2000US-0218290.  
XX

PR 14-AUG-2000; 2000US-0220963.  
XX

PR 14-AUG-2000; 2000US-0220964.  
XX

PR 14-AUG-2000; 2000US-0224518.  
XX

PR 14-AUG-2000; 2000US-0224519.  
XX

PR 14-AUG-2000; 2000US-0225213.  
XX

PR 14-AUG-2000; 2000US-0225266.  
XX

PR 14-AUG-2000; 2000US-0225267.  
XX

PR 14-AUG-2000; 2000US-0225268.  
XX

PR 14-AUG-2000; 2000US-0225270.  
XX

PR 14-AUG-2000; 2000US-0225447.  
XX

PR 14-AUG-2000; 2000US-0225757.  
XX

PR 14-AUG-2000; 2000US-0225758.  
XX

PR 14-AUG-2000; 2000US-0225759.  
XX

PR 18-AUG-2000; 2000US-0226279.  
XX

PR 22-AUG-2000; 2000US-0226681.  
XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 17:55:23 ; Search time 80 Seconds  
(without alignments)  
426.578 Million cell updates/sec

Title: US-09-936-680-2  
Perfect score: 1124  
Sequence: 1 MPAPNRLPLASLVLYWVS.....SDYLAIPSENKENSAPVVEE 215

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1124	100.0	215	21	Human beta3 subunit
2	1124	100.0	215	22	Human novel sodium
3	1124	100.0	215	23	Human signal trans
4	1105	98.3	215	21	Rat beta3 subunit
5	1105	98.3	215	22	Rat sodium channel
6	1024	91.1	195	22	Human protein SEQ
7	986	87.7	191	22	Rat sodium channel
8	942.5	83.9	369	22	Novel human diaph
9	843	75.0	159	21	Human beta3 subunit

10	832	74.0	159	21	AA336021	Rat beta3 subunit
11	694	61.7	1176	22	ABG22576	Novel human diaph
12	477	42.4	218	22	AA350260	Rabbit sodium chan
13	472	42.0	218	21	AA336031	Rat beta1 subunit
14	472	42.0	218	22	AA336031	Rat sodium channel
15	471	41.9	218	24	AB822771	Human voltage-gate
16	408.5	36.3	206	22	ABG22575	Novel human diaph
17	258.5	23.0	268	23	AAE18634	Human beta1A sodiu
18	246	21.9	272	22	AAE23320	Sodium channel bet
19	246	21.9	272	22	AB20371	Human voltage-gate
20	234	20.8	74	22	AB11698	Human protein SEQ
21	234	20.8	74	22	AA60196	Human beta3 subunit
22	207	18.4	39	21	AA336015	Rat beta3 subunit
23	196	17.4	39	21	AA336016	Human myelin prote
24	165	14.7	34	21	AA336029	Human beta3 subunit
25	163.5	14.5	258	23	AA52696	Human myelin prote
26	161	14.3	34	21	AA336034	Rat beta3 subunit
27	160.5	14.3	235	22	AA72879	Human PRO7425 prot
28	160.5	14.3	235	22	AA60121	Human transport pr
29	160.5	14.3	235	23	AA83710	Human PRO protein
30	160.5	14.3	235	24	ABG74763	Human PRO7425 prot
31	160.5	14.3	246	23	ABG79170	Human myelin-like
32	149	13.3	248	18	AA94567	Mouse major periph
33	147	13.1	215	20	AA41673	Human channel-rela
34	147	13.1	215	21	AA33448	Human PRO1192 prot
35	147	13.1	215	21	AA66752	Membrane-bound pro
36	147	13.1	215	22	AA12415	Human PRO1192 poly
37	147	13.1	215	22	AA65275	Human angiogenesis
38	147	13.1	215	23	AB89545	Human PRO1192 prot
39	147	13.1	215	23	AB84909	Human PRO protein
40	147	13.1	215	23	AA83675	Human PRO polypept
41	147	13.1	215	24	ABU66813	Human secreted/tra
42	147	13.1	215	24	ABU67089	Human secreted and
43	147	13.1	215	24	ABU59894	Novel human secret
44	147	13.1	215	24	ABU59168	Human secreted/tra
45	147	13.1	215	24	ABU59315	

## ALIGNMENTS

RESULT 1  
ID AAB36002 standard; protein; 215 AA.

XX AAB36002;

XX AC

XX 15-FEB-2001 (first entry)

XX Human beta3 subunit.

XX Human; beta sub-unit; beta3; analgesic; anticonvulsant;

XX cerebroprotective; vasotropic; cardiant; nootropic; cytostatic;

XX dermatological; gene therapy; voltage-gated sodium channel; pain;

XX epilepsy; stroke; ischaemia; heart disease; Jacobson Syndrome;

XX familial nonchronaffin paraganglioma; phenylketonuria;

XX Charcot Marie Tooth disease.

XX Homo sapiens.

XX WO2000063367-A1.

XX 26-OCT-2000.

XX 24-FEB-2000; 2000WO-EP01783.

XX 15-APR-1999; 99US-0129473.

XX (WARN ) WARNER LAMBERT CO.

XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Cox P, Dixon A, Jackson A, Morgan K;

DR WPI: 2000-665241/64.  
 DR N-PSDB; AAC67837.  
 XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
 PT sodium channel, and their corresponding polypeptides, useful for  
 PT detecting and treating sodium channel-associated conditions, e.g. pain,  
 PT epilepsy and stroke -  
 XX Claim 34; Fig 4; 88pp; English.  
 XX The present sequence is given in the claims of a specification  
 CC relating to a novel family of beta sub-unit proteins from a  
 CC voltage-gated sodium channel. Human and rat beta sub-units, which  
 CC have been collectively identified as beta3, have been isolated.  
 CC The polynucleotides and polypeptides are useful for screening for  
 CC agonists and antagonists of sodium channels. The agonists, antagonists,  
 CC proteins and nucleic acids may be used diagnosing of treating diseases  
 CC or conditions associated with voltage-gated sodium channels, e.g. pain,  
 CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial  
 CC Nonchromaffin paraganglioma, Phenylketonuria and Charcot Marie Tooth  
 CC disease.  
 XX Sequence 215 AA;  
 SQ

Query Match 100.0%; Score 1124; DB 21; Length 215;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-97;  
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAFNRLFLASLVLIIYVSVCFVPCVVEVPSETEAVQGNPKMLRCISCMKREVEATTVV 60  
 DB 1 MPAFNRLFLASLVLIIYVSVCFVPCVVEVPSETEAVQGNPKMLRCISCMKREVEATTVV 60  
 QY 61 EMFYRPEGKDFLIYRNHGOEVSPFQGRLOWNGSKDQVSTVNLVTLNDSGLYTC 120  
 DB 61 EMFYRPEGKDFLIYRNHGOEVSPFQGRLOWNGSKDQVSTVNLVTLNDSGLYTC 120  
 QY 121 NVSRPEFPAHPPFVKTRLLIPLRVTEEAGEDFTSVVSEIMMYILLVFLTLWLLIEMIYC 180  
 DB 121 NVSRPEFPAHPPFVKTRLLIPLRVTEEAGEDFTSVVSEIMMYILLVFLTLWLLIEMIYC 180  
 QY 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
 DB 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

RESULT 2  
 AAB85206  
 ID AAB85206 standard; Protein; 215 AA.  
 XX AAB85206;  
 XX 07-SEP-2001 (first entry)  
 XX Human novel sodium channel beta1-like subunit.  
 XX Sodium channel; sensory neurone specific channel; beta1-like subunit;  
 KW SNS; therapeutic; pain; analgesic.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Domain 38..122 "immunoglobulin domain"  
 FT Domain 157..176  
 FT Domain /note= "transmembrane domain"  
 XX WO200144293-A2.  
 XX 21-JUN-2001.  
 XX 14-DEC-2000; 2000WO-GB04802.  
 XX 17-DEC-1999; 99GB-0029970.

XX (GLAXO) GLAXO GROUP LTD.  
 XX Plumptre M, Powell AJ, Sanseau P;  
 XX WPI; 2001-398129/42.  
 DR N-PSDB; AAF84146.  
 XX Novel sub-unit for voltage-gated sodium channel proteins for producing  
 PT agents useful for treating pain -  
 XX Claim 1; Fig 2; 31pp; English.  
 XX The invention provides a novel beta1-like sub-unit for voltage-gated  
 CC sodium ion channel polypeptide, specifically a sensory neurone specific  
 CC channel (SNS) subunit. The novel beta1-like subunit is useful for  
 CC producing a therapeutic agent which is useful treating pain in a patient.  
 CC The subunit can be expressed by standard recombinant methodology. The  
 CC present sequence represents a human novel sodium channel beta1-like  
 CC subunit.  
 XX Sequence 215 AA;  
 SQ

Query Match 100.0%; Score 1124; DB 22; Length 215;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-97;  
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAFNRLFLASLVLIIYVSVCFVPCVVEVPSETEAVQGNPKMLRCISCMKREVEATTVV 60  
 DB 1 MPAFNRLFLASLVLIIYVSVCFVPCVVEVPSETEAVQGNPKMLRCISCMKREVEATTVV 60  
 QY 61 EMFYRPEGKDFLIYRNHGOEVSPFQGRLOWNGSKDQVSTVNLVTLNDSGLYTC 120  
 DB 61 EMFYRPEGKDFLIYRNHGOEVSPFQGRLOWNGSKDQVSTVNLVTLNDSGLYTC 120  
 QY 121 NVSRPEFPAHPPFVKTRLLIPLRVTEEAGEDFTSVVSEIMMYILLVFLTLWLLIEMIYC 180  
 DB 121 NVSRPEFPAHPPFVKTRLLIPLRVTEEAGEDFTSVVSEIMMYILLVFLTLWLLIEMIYC 180  
 QY 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
 DB 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

RESULT 3  
 ABB05689  
 ID ABB05689 standard; Protein; 215 AA.  
 XX ABB05689;  
 XX 30-APR-2002 (first entry)  
 XX Human signal transduction protein clone amy2\_2f18.  
 XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;  
 KW gene therapy.  
 XX Homo sapiens.  
 XX WO200198454-A2.  
 XX 27-DEC-2001.  
 XX 25-APR-2001; 2001WO-IB02050.  
 XX 25-APR-2000; 2000US-199380P.  
 XX (GEHU-) GERMAN HUMAN GENOME PROJECT.  
 XX Wiemann S;  
 XX WPI; 2002-055860/07.  
 DR N-PSDB; ABA93727.

XX Human cDNA sequences and clones derived from human fetal brain, fetal  
PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic  
PT screening and therapy -

XX Claim 1; Page 176; 61pp; English.

XX The present invention describes assemblages and computer readable media  
CC comprising novel human cDNA sequences and clones derived from human  
CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA  
CC libraries. ABA92702 to ABA93766 represent human cDNA sequences from the  
CC present invention which encode the proteins given in ABA95662 to  
CC ABA95729. The human cDNA sequences and clones can be used in gene  
CC therapy. The clones may be used in a variety of applications, for  
CC example they may be used in profiling assays, for providing large arrays  
CC of human genetic material for implementing large-scale screening  
CC strategies and for treating diseases via gene therapy procedures.

XX Sequence 215 AA;

Query Match 100.0%; Score 1124; DB 23; Length 215;  
Best Local Similarity 100.0%; Pred. No. 1.4e-97;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAFNRLFPPLASLVLIYVWVCPVCEVPSSTEAQGNPKLRCISCMKREVEATTVV 60  
DB |||||  
1 MPAFNRLFPPLASLVLIYVWVCPVCEVPSSTEAQGNPKLRCISCMKREVEATTVV 60  
QY 61 EWFYRPEGGKDFLIYEYRNGHQEVESPFQGRLOWNGSKLDQVSIITLVNLTNDGSLYTC 120  
DB |||||  
61 EWFYRPEGGKDFLIYEYRNGHQEVESPFQGRLOWNGSKLDQVSIITLVNLTNDGSLYTC 120  
QY 121 NVSRFEFEFAHRPFVKTTRLIPLRVTEEAGEDFTSVVSEIMYILLVFLTLMLLIEMIYC 180  
DB |||||  
121 NVSRFEFEFAHRPFVKTTRLIPLRVTEEAGEDFTSVVSEIMYILLVFLTLMLLIEMIYC 180  
QY 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
DB |||||  
181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

RESULT 4

AAB36001  
ID AAB36001 standard; protein; 215 AA.

AC AAB36001;

DT 15-FEB-2001 (first entry)

DE Rat beta3 subunit.

KW Rat; beta sub-unit; beta3; analgesic; anticonvulsant; cerebroprotective;  
KW vasotrophic; cardiac; nootropic; cytostatic; dermatological;  
KW gene therapy; voltage-gated sodium channel; pain; epilepsy; stroke;  
KW ischaemia; heart disease; Jacobsen Syndrome;  
KW familial nonchromaffin paraganglioma; phenylketonuria;  
KW Charcot Marie Tooth disease.

OS Rattus sp.

PN WO200063367-A1.

PD 26-OCT-2000.

PF 24-FEB-2000; 2000WO-EF01783.

PR 15-APR-1999; 99US-0129473.

PA (WARN ) WARNER LAMBERT CO.

XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

PI Cox P, Dixon A, Jackson A, Morgan K,

XX

DR WPI; 2000-665241/64.  
DR N-PSDB; AAC67836.

XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
PT sodium channel, and their corresponding polypeptides, useful for  
PT detecting and treating sodium channel-associated conditions, e.g. pain,  
PT epilepsy and stroke -

XX Claim 33; Fig 4; 88pp; English.

XX The present sequence is given in the claims of a specification  
CC relating to a novel family of beta sub-unit proteins from a  
CC voltage-gated sodium channel. Human and rat beta sub-units, which  
CC have been collectively identified as beta3, have been isolated.  
CC The polynucleotides and polypeptides are useful for screening for  
CC agonists and antagonists of sodium channels. The agonists, antagonists,  
CC proteins and nucleic acids may be used diagnosing of treating diseases  
CC or conditions associated with voltage-gated sodium channels, e.g. pain,  
CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial  
CC Nonchromaffin Paraganglioma, Phenylketonuria and Charcot Marie Tooth  
CC disease.

XX Sequence 215 AA;

Query Match 98.3%; Score 1105; DB 21; Length 215;  
Best Local Similarity 98.1%; Pred. No. 8.4e-96;  
Matches 211; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MPAFNRLFPPLASLVLIYVWVCPVCEVPSSTEAQGNPKLRCISCMKREVEATTVV 60  
DB |||||  
1 MPAFNRLFPPLASLVLIYVWVCPVCEVPSSTEAQGNPKLRCISCMKREVEATTVV 60  
QY 61 EWFYRPEGGKDFLIYEYRNGHQEVESPFQGRLOWNGSKLDQVSIITLVNLTNDGSLYTC 120  
DB |||||  
61 EWFYRPEGGKDFLIYEYRNGHQEVESPFQGRLOWNGSKLDQVSIITLVNLTNDGSLYTC 120  
QY 121 NVSRFEFEFAHRPFVKTTRLIPLRVTEEAGEDFTSVVSEIMYILLVFLTLMLLIEMIYC 180  
DB |||||  
121 NVSRFEFEFAHRPFVKTTRLIPLRVTEEAGEDFTSVVSEIMYILLVFLTLMLLIEMIYC 180  
QY 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
DB |||||  
181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

RESULT 5

AAB50243

ID AAB50243 standard; protein; 215 AA.

AC AAB50243;

DT 13-MAR-2001 (first entry)

DE Rat sodium channel beta3 protein Alrxa94h5.

KW Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;  
KW neurodegenerative disorder; mood disorder; muscle contraction.

OS Rattus sp.

PN WO200069912-A1.

PD 23-NOV-2000.

PF 12-MAY-2000; 2000WO-US13144.

PR 14-MAY-1999; 99US-0134198.

PA (MILL-) MILLENNIUM PHARM INC.

PI Curtis RAJ;

DR WPI; 2001-122743/13.

DR N-PSDB; AAC90600.

XX New rat sodium channel beta-3 subunit gene isolated from a rat dorsal

PT root ganglion cDNA library for use in chromosome mapping, forensic

PT medicine, monitoring clinical trials and therapeutics -

XX

XX Claim 9; Fig 2; 145pp; English.

XX

CC The present invention provides the protein and coding sequences of the

CC rat sodium channel beta3 protein, designated Alrxa94h5. This protein is

CC involved in the generation of pain and other sensory or perceptible nerve

CC impulses, in the establishment and endurance of mood, neurodegenerative

CC and sleep disorders, and in the control of muscle contraction, including

CC movements such as the heartbeat, digestion and vascular tone. The

CC sequences can be used in predictive medicine, screening and diagnostic

CC assays, and in pharmacogenomics.

XX

XX Sequence 215 AA;

SQ

Query Match 98.3%; Score 1105; DB 22; Length 215;

Best Local Similarity 98.1%; Pred. No. 8.4e-96;

Matches 211; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MPANRLLPLASLVLIYVSVCPVCEVPSSEAVQGNPKMLRCISCMKREVEATTVV 60

DB 1 MPANRLLPLASLVLIYVSVCPVCEVPSSEAVQGNPKMLRCISCMKREVEATTVV 60

QY 61 EWFYRPEGKDFLIYVYRNGHVEVSPFQGRLOWNGSKDLQDVSIITLVNLTNDGSLYTC 120

DB 61 EWFYRPEGKDFLIYVYRNGHVEVSPFQGRLOWNGSKDLQDVSIITLVNLTNDGSLYTC 120

QY 121 NVSRREFEFAHRPFVKTTLRIPLRVTEAGEDFTSVSEIMYILLVFLTLWLLIEMIYC 180

DB 121 NVSRREFEFAHRPFVKTTLRIPLRVTEAGEDFTSVSEIMYILLVFLTLWLLIEMIYC 180

QY 181 YRKVSKAEAAQENASDYLAIPESENKENSAPVVEE 215

DB 181 YRKVSKAEAAQENASDYLAIPESENKENSAPVVEE 215

RESULT 6

AAW79212

ID AAW79212 standard; Protein; 195 AA.

AC AAW79212;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human protein SEQ ID NO 1874.

XX

Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

XX

OS Homo sapiens.

XX

XX WO200157190-A2.

XX

XX 09-AUG-2001.

XX

XX 05-FEB-2001; 2001WO-US04098.

XX

XX 03-FEB-2000; 2000US-0496914.

XX

XX 27-APR-2000; 2000US-0560875.

XX

XX 20-JUN-2000; 2000US-0598075.

XX

XX 19-JUL-2000; 2000US-0620325.

XX

XX 01-SEP-2000; 2000US-0654936.

XX

XX 15-SEP-2000; 2000US-0663561.

XX

XX 20-OCT-2000; 2000US-0693325.

XX

XX 30-NOV-2000; 2000US-0728422.

XX

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

XX WPI; 2001-476283/51.

DR N-PSDB; AAK52345.

XX

XX Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

XX

XX Claim 20; Page 4260; 6221pp; English.

PS

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

CC (AAM80020) are omitted as the relevant pages from the sequence listing

CC were missing at the time of publication.

XX

SQ Sequence 195 AA;

Query Match 91.1%; Score 1024; DB 22; Length 195;

Best Local Similarity 100.0%; Pred. No. 3e-88;

Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPANRLLPLASLVLIYVSVCPVCEVPSSEAVQGNPKMLRCISCMKREVEATTVV 60

DB 1 MPANRLLPLASLVLIYVSVCPVCEVPSSEAVQGNPKMLRCISCMKREVEATTVV 60

QY 61 EWFYRPEGKDFLIYVYRNGHVEVSPFQGRLOWNGSKDLQDVSIITLVNLTNDGSLYTC 120

DB 61 EWFYRPEGKDFLIYVYRNGHVEVSPFQGRLOWNGSKDLQDVSIITLVNLTNDGSLYTC 120

QY 121 NVSRREFEFAHRPFVKTTLRIPLRVTEAGEDFTSVSEIMYILLVFLTLWLLIEMIYC 180

DB 121 NVSRREFEFAHRPFVKTTLRIPLRVTEAGEDFTSVSEIMYILLVFLTLWLLIEMIYC 180

QY 181 YRKVSKAEAAQENASDYLAIPESENKENSAPVVEE 195

DB 181 YRKVSKAEAAQENASDYLAIPESENKENSAPVVEE 195

RESULT 7

AAW50245

ID AAW50245 standard; Protein; 191 AA.

XX

XX AAW50245;

XX

XX 13-MAR-2001 (first entry)

XX

XX Rat sodium channel beta3 protein Alrxa94h5 mature protein.

XX

XX Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;

KW neurodegenerative disorder; mood disorder; muscle contraction.

XX

XX Rattus sp.

XX

XX WO200069912-A1.

XX

XX 23-NOV-2000.

XX

XX 12-MAY-2000; 2000WO-US13144.

XX

XX 14-MAY-1999; 99US-0134198.

XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Curtis RAJ;  
XX XX  
DR WPI; 2001-122743/13.  
DR N-PSDB; AAC90600.  
XX XX  
PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal  
PT root ganglion cDNA library for use in chromosome mapping, forensic  
PT medicine, monitoring clinical trials and therapeutics -  
XX XX  
PS Claim 9; Fig 2; 145pp; English.  
XX XX  
CC The present invention provides the protein and coding sequences of the  
CC rat sodium channel beta3 protein, designated Alrx94h5. This protein is  
CC involved in the generation of pain and other sensory or perceptible nerve  
CC impulses, in the establishment and endurance of mood, neurodegenerative  
CC and sleep disorders, and in the control of muscle contraction, including  
CC movements such as the heartbeat, digestion and vascular tone. The  
CC sequences can be used in predictive medicine, screening and diagnostic  
CC assays, and in pharmacogenomics.  
XX XX  
SQ Sequence 191 AA;  
Query Match 87.7%; Score 986; DB 22; Length 191;  
Best Local Similarity 99.0%; Pred. No. 1.1e-84;  
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 25 VCVEPSETEAVQGNPKLRCISCKREVEATTVEWYRPEGKDFLIYEYRNGHQEV 84  
Db 1 VCVEPSETEAVQGNPKLRCISCKREVEATTVEWYRPEGKDFLIYEYRNGHQEV 60  
QY 85 ESPFQGRLOMNGSKDLQDVSTVLNVTINDSGLYTCNVSRFEFEARHPFVKTTTLPLR 144  
Db 61 ESPFQGRLOMNGSKDLQDVSTVLNVTINDSGLYTCNVSRFEFEARHPFVKTTTLPLR 120  
QY 145 VTEAGDEFTSVSEIMMYILLVFLTLWLLIEMIVCYRKVSKABEAAQENASDYLAIPSE 204  
Db 121 VTEAGDEFTSVSEIMMYILLVFLTLWLLIEMIVCYRKVSKABEAAQENASDYLAIPSE 180  
QY 205 NKENSVPVEE 215  
Db 181 NKENSVPVEE 191  
RESULT 8  
ABG22577  
ID ABG22577 standard; protein; 369 AA.  
XX AC ABG22577;  
XX DT 18-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #22568.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX XX 30-MAR-2001; 2001WO-US08631.  
XX PR 31-MAR-2000; 2000US-0540217.  
XX PR 23-AUG-2000; 2000US-0649167.  
XX XX (HYSE-) HYSEQ INC.  
XX PA Drmanac RT, Liu C, Tang YT;  
XX PI

XX DR WPI; 2001-639362/73.  
XX N-PSDB; AAS86764.  
XX XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX XX  
PS Claim 20; SEQ ID No 52936; 103pp; English.  
XX XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX XX  
SQ Sequence 369 AA;  
Query Match 83.9%; Score 942.5; DB 22; Length 369;  
Best Local Similarity 93.9%; Pred. No. 3.1e-80;  
Matches 184; Conservative 2; Mismatches 9; Indels 1; Gaps 1;  
QY 1 MPAFNRLFLPLASLVLIYVSVCPFPVGVSETEA-VQGNPKLRCISCKREVEATTV 59  
Db 127 MPAFNRLFLPLASLVLIYVSVCPFPVGVSETEA-VQGNPKLRCISCKREVEATTV 186  
QY 60 VEMFYRPEGKDFLIYEYRNGHQEVSPFQGRLOMNGSKDLQDVSTVLNVTINDSGLYT 119  
Db 187 VEMFYRPEGKDFLIYEYRNGHQEVSPFQGRLOMNGSKDLQDVSTVLNVTINDSGLYT 246  
QY 120 CNVSRFEFEARHPFVKTTTLPLRVTTEAGDEFTSVSEIMMYILLVFLTLWLLIEMIV 179  
Db 247 CNVSRFEFEARHPFVKTTTLPLRVTTEAGDEFTSVSEIMMYILLVFLTLWLLIEMIV 306  
QY 180 CYRKVSKABEAAQENA 195  
Db 307 CYRKVSKABEAAQENA 322  
RESULT 9  
AAB36020  
ID AAB36020 standard; protein; 159 AA.  
XX AC AAB36020;  
XX DT 15-FEB-2001 (first entry)  
XX DE Human beta3 subunit peptide, SEQ ID NO: 22.  
XX KW Human; beta sub-unit; beta3; analgesic; anticonvulsant;  
XX cerebroprotective; vasotropic; cadiant; nootropic; cytosstatic;  
XX dermatological; gene therapy; voltage-gated sodium channel; pain;  
XX epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;  
XX familial nonchromaffin paraganglioma; phenylketonuria;  
XX Charcot Marie Tooth disease.  
XX XX

```
OS Homo sapiens.
XX WO200063367-A1.
XX PD 26-OCT-2000.
XX PF 24-FEB-2000; 2000WO-EP01783.
XX PR 15-APR-1999; 99US-0129473.
XX PA (WARN ) WARNER LAMBERT CO.
XX PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
XX PI Cox P, Dixon A, Jackson A, Morgan K;
XX WPI; 2000-665241/64.
XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated
PT sodium channel, and their corresponding polypeptides, useful for
PT detecting and treating sodium channel-associated conditions, e.g. pain,
PT epilepsy and stroke -
XX Claim 36; Page 76; 88pp; English.
XX The present sequence is given in the claims of a specification
CC relating to a novel family of beta sub-unit proteins from a
CC voltage-gated sodium channel. Human and rat beta sub-units, which
CC have been collectively identified as beta3, have been isolated.
CC The polynucleotides and polypeptides are useful for screening for
CC agonists and antagonists of sodium channels. The agonists, antagonists,
CC proteins and nucleic acids may be used diagnosing of treating diseases
CC or conditions associated with voltage-gated sodium channels, e.g. pain,
CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial
CC Nonchromaffin Paraganglioma, Phenylketonuria and Charcot Marie Tooth
CC disease.
XX SQ Sequence 159 AA;
Query Match 75.0%; Score 843; DB 21; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.4e-71;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPAFNRLLPLASLLIYVSVCFVCPVEPSETEAVQGNPMKLCISCMKREEVEATTVV 60
DB 1 MPAFNRLLPLASLLIYVSVCFVCPVEPSETEAVQGNPMKLCISCMKREEVEATTVV 60
QY 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITVLTNDSGLYTC 120
DB 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITVLTNDSGLYTC 120
QY 121 NVSREFEFEAHRPFVKTTRLIPLRVTEAGEDFTSVVSE 159
DB 121 NVSREFEFEAHRPFVKTTRLIPLRVTEAGEDFTSVVSE 159
RESULT 10
AAB36021
ID AAB36021 standard; protein; 159 AA.
XX AC AAB36021;
XX DT 15-FEB-2001 (first entry)
XX DE Rat beta3 subunit peptide, SEQ ID NO: 23.
XX Rat; beta sub-unit; beta3; analgesic; anticonvulsant;
KW cerebroprotective; vasotropic; cardiant; nootropic; cytostatic;
KW dermatological; gene therapy; voltage-gated sodium channel; pain;
KW epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;
KW familial nonchromaffin paraganglioma; phenylketonuria;
KW Charcot Marie Tooth disease.
XX Rattus sp.
OS
```

```
XX WO200063367-A1.
XX PD 26-OCT-2000.
XX PF 24-FEB-2000; 2000WO-EP01783.
XX PR 15-APR-1999; 99US-0129473.
XX PA (WARN ) WARNER LAMBERT CO.
XX PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
XX PI Cox P, Dixon A, Jackson A, Morgan K;
XX WPI; 2000-665241/64.
XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated
PT sodium channel, and their corresponding polypeptides, useful for
PT detecting and treating sodium channel-associated conditions, e.g. pain,
PT epilepsy and stroke -
XX Claim 36; Page 76-77; 88pp; English.
XX The present sequence is given in the claims of a specification
CC relating to a novel family of beta sub-unit proteins from a
CC voltage-gated sodium channel. Human and rat beta sub-units, which
CC have been collectively identified as beta3, have been isolated.
CC The polynucleotides and polypeptides are useful for screening for
CC agonists and antagonists of sodium channels. The agonists, antagonists,
CC proteins and nucleic acids may be used diagnosing of treating diseases
CC or conditions associated with voltage-gated sodium channels, e.g. pain,
CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial
CC Nonchromaffin Paraganglioma, Phenylketonuria and Charcot Marie Tooth
CC disease.
XX SQ Sequence 159 AA;
Query Match 74.0%; Score 832; DB 21; Length 159;
Best Local Similarity 98.7%; Pred. No. 2.6e-70;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MPAFNRLLPLASLLIYVSVCFVCPVEPSETEAVQGNPMKLCISCMKREEVEATTVV 60
DB 1 MPAFNRLLPLASLLIYVSVCFVCPVEPSETEAVQGNPMKLCISCMKREEVEATTVV 60
QY 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITVLTNDSGLYTC 120
DB 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITVLTNDSGLYTC 120
QY 121 NVSREFEFEAHRPFVKTTRLIPLRVTEAGEDFTSVVSE 159
DB 121 NVSREFEFEAHRPFVKTTRLIPLRVTEAGEDFTSVVSE 159
RESULT 11
ABG22576
ID ABG22576 standard; Protein; 1176 AA.
XX AC ABG22576;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #22567.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX PD 11-OCT-2001.
XX
```



PN WO200063367-A1.  
 XX 26-OCT-2000.  
 XX 24-FEB-2000; 2000WO-EP01783.  
 XX 15-APR-1999; 99US-0129473.  
 XX (WARN ) WARNER LAMBERT CO.  
 XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.  
 XX Cox P, Dixon A, Jackson A, Morgan K;  
 XX WPI; 2000-665241/64.  
 XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
 PT sodium channel, and their corresponding polypeptides, useful for  
 PT detecting and treating sodium channel-associated conditions, e.g. pain,  
 PT epilepsy and stroke -  
 XX Disclosures; Fig 4; 88pp; English.  
 XX The present sequence is given in a specification relating to a novel  
 CC family of beta sub-unit proteins from a voltage-gated sodium channel.  
 CC Human and rat beta sub-units, which have been collectively identified  
 CC as beta3, have been isolated. The polynucleotides and polypeptides are  
 CC useful for screening for agonists and antagonists of sodium channels.  
 CC The agonists, antagonists, proteins and nucleic acids may be used  
 CC in the diagnosis of treating diseases or conditions associated with  
 CC voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischaemia,  
 CC heart disease, Jacobsen Syndrome, Familial Nonchronaffin  
 CC Paraganglioma, Phenylketonuria and Charcot Marie Tooth disease.  
 XX Sequence 218 AA;  
 SQ Query Match 42.0%; Score 472; DB 21; Length 218;  
 Best Local Similarity 49.1%; Pred. No. 2.7e-36;  
 Matches 105; Conservative 33; Mismatches 68; Indels 8; Gaps 5;  
 QY 10 LASLVLYVSVCFVCPVEPSETEAVQGNPKLCISCKMKEEVEATTVWVFRPEGG 69  
 DB 5 LALVVGAVLVSSAWGCGVEVDSETEAVYGMTFKILCISCKRSETTAETFTTFRQKT 64  
 QY 70 KDFL-IYEVRNGHQEVESP--FQGRLOWNGS---KDLQDVSTVLNVLNDSGLTYCNVS 123  
 DB 65 EEFVKILRYENEVLQLEDEDERFEGRVVNGSRGTQDLQDLSIFITNVTYHSGDYECVY 124  
 QY 124 REFEFAHRFPVKTTLPLRVTEBAGEDFTSVSEIMMYILLVPLTLWLLIEMTYCYRK 183  
 DB 125 RLAFPDNYEHNTSVVKKHLEVDVKANRDMASIVSEIMMYILLVPLTLWLVAEVYCYK 184  
 QY 184 VSKA-EEAQAENASDYLAIPTSENKEN-SAVPVVEE 215  
 DB 185 IAAATEAQAENASEYLAIPTSEKENCCTGVQVAE 218  
 RESULT 14  
 AAB62326  
 ID AAB62326 standard; Protein; 218 AA.  
 XX AAB62326;  
 XX 29-JUN-2001 (first entry)  
 XX Rat sodium channel beta-1 subunit amino acid sequence.  
 XX Sodium channel; modulator; sodium channel beta-1A subunit; pain; rat;  
 XX sodium channel beta-1 subunit; analgesic; neuroprotective.  
 XX Rattus sp.  
 XX WO200123570-A2.  
 XX

PD 05-APR-2001.  
 XX 29-SEP-2000; 2000WO-US27034.  
 XX 30-SEP-1999; 99US-0156837.  
 XX (ORTH ) ORTHO-MCNEIL PHARM INC.  
 XX D'Andrea M, Rogers KE;  
 XX WPI; 2001-281683/29.  
 XX Screening for sodium channel activity modulators, used to decrease  
 PT neuropathic pain, comprises contacting a candidate compound with a cell  
 PT expressing the channel -  
 XX Examples; Fig 1B; 124pp; English.  
 XX The invention relates to a method of screening for a modulator of sodium  
 CC channel activity that comprises contacting a candidate modulator with a  
 CC cell co-expressing a sodium channel beta-1A subunit with a sodium channel  
 CC alpha subunit, and determining the effect of the candidate modulator on  
 CC the sodium channel function in the cell. The method is useful for  
 CC identifying sodium channel activity modulators, preferably causing  
 CC decreased beta 1A subunit expression. The modulators can be used to  
 CC decrease neuropathic pain, and to decrease the number of febrile seizures  
 CC in an individual. The present sequence represents a rat sodium channel  
 CC beta-1 subunit, used in homology studies with the beta-1A subunit.  
 XX Sequence 218 AA;  
 SQ Query Match 42.0%; Score 472; DB 22; Length 218;  
 Best Local Similarity 49.1%; Pred. No. 2.7e-36;  
 Matches 105; Conservative 33; Mismatches 68; Indels 8; Gaps 5;  
 QY 10 LASLVLYVSVCFVCPVEPSETEAVQGNPKLCISCKMKEEVEATTVWVFRPEGG 69  
 DB 5 LALVVGAVLVSSAWGCGVEVDSETEAVYGMTFKILCISCKRSETTAETFTTFRQKT 64  
 QY 70 KDFL-IYEVRNGHQEVESP--FQGRLOWNGS---KDLQDVSTVLNVLNDSGLTYCNVS 123  
 DB 65 EEFVKILRYENEVLQLEDEDERFEGRVVNGSRGTQDLQDLSIFITNVTYHSGDYECVY 124  
 QY 124 REFEFAHRFPVKTTLPLRVTEBAGEDFTSVSEIMMYILLVPLTLWLLIEMTYCYRK 183  
 DB 125 RLAFPDNYEHNTSVVKKHLEVDVKANRDMASIVSEIMMYILLVPLTLWLVAEVYCYK 184  
 QY 184 VSKA-EEAQAENASDYLAIPTSENKEN-SAVPVVEE 215  
 DB 185 IAAATEAQAENASEYLAIPTSEKENCCTGVQVAE 218  
 RESULT 15  
 AAB62771  
 ID AAB62771 standard; Protein; 218 AA.  
 XX AAB62771;  
 XX 18-MAR-2003 (first entry)  
 XX Human voltage-gated sodium channel mutant beta-1 subunit.  
 XX Sodium channel; mutation; epilepsy; anticonvulsant; mutant; human;  
 XX SCN1B; gene; ds.  
 XX Homo sapiens.  
 XX WO200290532-A1.  
 XX 14-NOV-2002.  
 XX 09-MAY-2002; 2002WO-AU00581.  
 XX



A:Reference number: A42737; MUID:92271207; PMID:1375395

A:Accession: A42737

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid; protein

A:Residues: 1-218 <ISO>

A:CROSS-references: GB:M91808; MID:9206864; PIDN:AAA88513.1; PID:9206865

A:Experimental source: brain

A>Note: sequence extracted from NCBI backbone (NCBIP:103298)

C:Keywords: transmembrane protein

Query Match 42.0%; Score 472; DB 2; Length 218;

Best Local Similarity 49.1%; Pred. No. 6.7e-33;

Matches 105; Conservative 33; Mismatches 68; Indels 8; Gaps 5;

QY 10 LASLVLIYWSVCFVCEVPESETEAVQGNPMKLRICISCMKREVEATTVVWFYRPEG 69

DB 5 LALVGVAVLVSSAGWCGVEVDSETEAVYGMTFKILCISCKRSETTAETFTFRQGT 64

QY 70 KDFL-IYERNHQHVESP--FQGRLOWNGS---KDLQDVSIIVNLVTLNDSGLYTCNVS 123

DB 65 EEFVKILRYENEVLQLEEDERPEGRVWVWNGSGRTKDLQDLGIFITNVTYHNSGDYECVY 124

QY 124 REFEPFAHRPFVKTTRLIPLRVTEAGEDFTSVVSEIMMYILLVFLTLMLLIEMICYRK 183

DB 125 RLLFDNYEHNTSVVKKIHLEVDKANKDMASIVSEIMMYILLVFLTLMLVAEMICYRK 184

QY 184 VSKA-EAAQENASDYLAIPSENKEN-SAVPVEE 215

DB 185 IAAATEAAQENASVLAITSEKNCCTGVQVAE 218

RESULT 3

AS5734

sodium channel, voltage-gated, beta-1 chain precursor - human

C:Species: Homo sapiens (man)

C>Date: 18-Aug-1995 #sequence revision 18-Aug-1995 #text\_change 08-Oct-1999

C:Accession: A55734; A53408; I54354

R:Mapita, N.; Sloan-Brown, K.; Weghuis, D.O.; Ropers, H.H.; George Jr., A.L.

Genomics 23, 628-634, 1994

A:Title: Genomic organization and chromosomal assignment of the human voltage-gated Na(+)

A:Reference number: A55734; MUID:95154833; PMID:7851891

A:Accession: A55734

A:Molecule type: DNA

A:Residues: 1-218 <MAK1>

A:CROSS-references: GB:U12189

R:Mapita, N.; Bennett Jr., P.B.; George Jr., A.L.

J. Biol. Chem. 269, 7571-7578, 1994

A:Title: Voltage-gated Na(+) channel beta-1 subunit mRNA expressed in adult human skeletal

A:Reference number: A53408; MUID:94171787; PMID:8125980

A:Accession: A53408

A:Molecule type: mRNA

A:Residues: 1-218 <MAK2>

A:CROSS-references: GB:L16242; MID:9450602; PIDN:AAA61277.1; PID:9450603

R:McClatchey, A.L.; Cannon, S.C.; Slaugenhaupt, S.A.; Gusella, J.F.

Hum. Mol. Genet. 2, 745-749, 1993

A:Title: The cloning and expression of a sodium channel beta 1-subunit cDNA from human brain

A:Reference number: I54354; MUID:93357746; PMID:8394762

A:Accession: I54354

A:Molecule type: mRNA

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-218 <RES>

A:CROSS-references: GB:L10338; MID:9307414; PIDN:AAA60391.1; PID:9307415

A:Gene: GDB:SCN1B

A:Map position: 19q13.1-19q13.1

C:Keywords: glycoprotein; muscle; transmembrane protein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:161-182/Domain: transmembrane #status predicted <TMD>

F:93,110,114,135/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.9%; Score 471; DB 2; Length 218;

Best Local Similarity 49.5%; Pred. No. 8.1e-33;

QY 10 LASLVLIYWSVCFVCEVPESETEAVQGNPMKLRICISCMKREVEATTVVWFYRPEG 69

DB 5 LALVGVAVLVSSAGWCGVEVDSETEAVYGMTFKILCISCKRSETTAETFTFRQGT 64

QY 70 KDFL-IYERNHQHVESP--FQGRLOWNGS---KDLQDVSIIVNLVTLNDSGLYTCNVS 123

DB 65 EEFVKILRYENEVLQLEEDERPEGRVWVWNGSGRTKDLQDLGIFITNVTYHNSGDYECVY 124

QY 124 REFEPFAHRPFVKTTRLIPLRVTEAGEDFTSVVSEIMMYILLVFLTLMLLIEMICYRK 183

DB 125 RLLFDNYEHNTSVVKKIHLEVDKANKDMASIVSEIMMYILLVFLTLMLVAEMICYRK 184

QY 184 VSKA-EAAQENASDYLAIPSENKEN-SAVPVEE 215

DB 185 IAAATEAAQENASVLAITSEKNCCTGVQVAE 218

Matches 106; Conservative 31; Mismatches 69; Indels 8; Gaps 5;

QY 10 LASLVLIYWSVCFVCEVPESETEAVQGNPMKLRICISCMKREVEATTVVWFYRPEG 69

DB 5 LALVGVAVLVSSAGWCGVEVDSETEAVYGMTFKILCISCKRSETTAETFTFRQGT 64

QY 70 KDFL-IYERNHQHVESP--FQGRLOWNGS---KDLQDVSIIVNLVTLNDSGLYTCNVS 123

DB 65 EEFVKILRYENEVLQLEEDERPEGRVWVWNGSGRTKDLQDLGIFITNVTYHNSGDYECVY 124

QY 124 REFEPFAHRPFVKTTRLIPLRVTEAGEDFTSVVSEIMMYILLVFLTLMLLIEMICYRK 183

DB 125 RLLFDNYEHNTSVVKKIHLEVDKANKDMASIVSEIMMYILLVFLTLMLVAEMICYRK 184

QY 184 VSKA-EAAQENASDYLAIPSENKEN-SAVPVEE 215

DB 185 IAAATEAAQENASVLAITSEKNCCTGVQVAE 218

RESULT 4

MPETO

myelin P0 protein precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 18-Feb-1995

C:Accession: JQ6622; A22822

R:Lemke, G.; Lamm, E.; Patterson, J.

Neuron 1, 73-83, 1988

A:Title: Isolation and analysis of the gene encoding peripheral myelin protein zero.

A:Reference number: JQ6622; MUID:90166482; PMID:2483091

A:Accession: JQ6622

A:Molecule type: DNA

A:Residues: 1-248 <LEM>

R:Lemke, G.; Axel, R.

Cell 40, 501-508, 1985

A:Title: Isolation and sequence of a cDNA encoding the major structural protein of per

A:Reference number: A22822; MUID:85124601; PMID:2578885

A:Accession: A22822

A:Molecule type: mRNA

A:Residues: 1-248 <LEM2>

C:Comment: This protein is found only in peripheral nervous system Schwann cells.

C:Genetics:

A:Introns: 23/1; 78/3; 150/1; 195/2; 215/3

C:Superfamily: myelin P0 protein; immunoglobulin homology

C:Keywords: glycoprotein; myelin; Schwann cell; structural protein; transmembrane prote

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-248/Product: myelin P0 protein #status predicted <MAT>

F:30-153/Domain: extracellular #status predicted <EXD>

F:43-129/Domain: immunoglobulin homology <IMM>

F:154-179/Domain: transmembrane #status predicted <TMD>

F:180-248/Domain: intracellular #status predicted <IND>

F:122/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.6%; Score 175; DB 1; Length 248;

Best Local Similarity 28.1%; Pred. No. 1.4e-07;

Matches 59; Conservative 34; Mismatches 99; Indels 18; Gaps 7;

QY 11 ASLVLIYWSVCFVCEVPESETEAVQGNPMKLRICISCMKREVEATTVVWFYRPEG 70

DB 20 SSLVL---SPTLAIWYTDREVYGAAGVQVTLHC-SFMSSEWSDDISFTWRYQEGGR 74

QY 71 DFL-IYERNHQHVESP--FQGRLOWNGSKDLQDVSIIVNLVTLNDSGLYTCNVSREFE 127

DB 75 DAISIFHYAKGFYIDEVGTFKERTIQWVGDPDSWKDGSIVIHNLDSNGTGTCDVKSPDP 134

QY 128 FEHRPFVKTTRLIPLRVTEAGEDFTSVVSEIMMYILLVFLTLMLLIEMICYRK 186

DB 135 -----IVGKTSQVLYYFEKVPYRGVGVJGNAVIGILGVLLLLLLFLYRYCWLRRQA 188

QY 187 ABE---AAQENASDYLAIPSENKENSAVPV 213

DB 189 ALQRRLSAMEKGFHKSSKDSKRGRTFV 218

```
RESULT 5
JH0252
myelin P0 protein precursor - human
N;Alternate names: myelin protein zero
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 06-Dec-1996 #text_change 16-Jun-2000
C;Accession: JH0252; JN0704; F13978; I58118
R;Hayasaka, K.; Nanao, K.; Tahara, M.; Sato, M.; Uyemura, K.
Biochem. Biophys. Res. Commun. 180, 515-518, 1991
A;Title: Isolation and sequence determination of cDNA encoding the major structural protein of myelin P0
A;Reference number: JH0252; MUID:92062068; PMID:1719967
A;Accession: JH0252
A;Molecule type: mRNA
A;Residues: 1-248 <HAY>
A;Cross-references: GB:D10537; GB:D90501; NID:G220073; PID:BAA01395.1; PID:G220074
R;Hayasaka, K.; Ohnishi, A.; Takada, G.; Fukushima, Y.; Murai, Y.
Biochem. Biophys. Res. Commun. 194, 1317-1322, 1993
A;Title: Mutation of the myelin P0 gene in charcot-Marie-Tooth neuropathy type 1.
A;Reference number: JN0704; MUID:93356807; PMID:7688964
A;Accession: JN0704
A;Molecule type: mRNA
A;Residues: 1-248 <HA2>
R;Pham-Dinh, D.; Fourbil, Y.; Blanquet, F.; Mattei, M.G.; Roedel, N.; Latour, P.; Chazotte, M.
Hum. Mol. Genet. 2, 2051-2054, 1993
A;Title: The major peripheral myelin protein zero gene: structure and localization in the human genome
A;Reference number: I39378; MUID:94154677; PMID:7509228
A;Accession: I39378
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-248 <RES>
A;Cross-references: GB:I24893; NID:G454412; PID:AAA20656.1; PID:G529405
R;Kulkarni, T.; Bolhuis, P.A.; Wolterman, R.A.; Kemp, S.; te Nijenhuis, S.; Valentijn, L.
Nature Genet. 5, 35-39, 1993
A;Title: Deletion of the serine 14 codon from the major peripheral myelin protein P0 gene
A;Reference number: I58118; MUID:94035114; PMID:7693130
A;Accession: I58118
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-62, 64-115 <RE2>
A;Cross-references: GB:S66705; NID:G437048; PID:AAB28708.1; PID:G437049
A;Experimental source: disease-state mutant
A;Comment: This protein, a small integral membrane glycoprotein, is the most abundant protein in myelin.
A;Comment: This protein plays essential roles in both the elaboration and the subsequent myelination of axons.
C;Genetics:
A;Gene: GDB:MPZ; CMT1B; HMSNIB
A;Cross-references: GDB:125266; OMIM:159440
A;Map position: 1q22-1q23
A;Introns: 23/1; 78/3; 150/1; 195/2; 215/3
C;Superfamily: myelin P0 protein; immunoglobulin homology
C;Keywords: glycoprotein; membrane protein; myelin; Schwann cell; structural protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:30-248/Product: peripheral myelin #status predicted <PER>
F:43-129/Domain: immunoglobulin homology <IMM>
F:50-127/Disulfide bonds: #status predicted
F:122/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 14.5%; Score 163.5; DB 1; Length 248;
Best Local Similarity 28.3%; Pred. No. 1.3e-06;
Matches 54; Conservative 33; Mismatches 81; Indels 23; Gaps 7;

QY 11 ASLVLYVSVCFPCVVEPSTEAVQGNPKMLRCISCKREVEATTVVEWFFYRPEGK 70
DB 20 SSLVL-----SPAQAIVVYTDREHVGAVGSRVTLHC-SFWSSEWSDDISFTWRYQPEGGR 74

QY 71 DFL-IYEVNNGHQEVE--SPFGRLQWNGSKDLQDVSTVLNVLNDSGLYTCNVSRFFE 127
DB 75 DAISIFHYAKGPYIDEVGTGFKERIQWVGDPKWDGSIHNLDSYDNGTFTCDVKNPPD 134

QY 128 FEHRPFVKTTLRLPLRVTEAGEDFTSVSVSEIMMYILLVFLTLWLLIEMI-YCY----- 181
DB 135 -----IVGKTSQVTLVYFEKVPTRYGVVIGVGLGVLLVLLVLLVYVRYCWLRRQA 188

C;Keywords: glycoprotein; membrane protein; myelin; Schwann cell; structural protein
```

```
QY 182 ---RKVSKAE 189
DB 189 ALQRRLSAMEK 199

RESULT 6
I38053
myelin protein zero - human
C;Species: Homo sapiens (man)
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 21-Jan-2000
C;Accession: I38053; A49643; S43191
R;Rautenstrauss, B.; Nelles, E.; Grehl, H.; Pfeiffer, R.A.; Van Broeckhoven, C.
Hum. Mol. Genet. 3, 1701-1702, 1994
A;Title: Identification of a de novo insertional mutation in P0 in a patient with a Dejerine-Sottas syndrome
A;Reference number: I38053; MUID:95135435; PMID:7530550
A;Accession: I38053
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-251 <RES>
A;Cross-references: EMBL:Z31718; NID:G469516; PID:CNA83513.1; PID:G469517
R;Su, Y.; Brooks, D.G.; Li, L.; Lepercq, J.; Trotter, J.A.; Ravetch, J.V.; Lebo, R.V.
Proc. Natl. Acad. Sci. U.S.A. 90, 10856-10860, 1993
A;Title: Myelin protein zero gene mutated in Charcot-Marie-Tooth type 1B patients.
A;Reference number: A49643; MUID:94068501; PMID:7504284
A;Accession: A49643
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 50-105 <SUL>
A;Experimental source: peripheral blood
A;Note: sequence extracted from NCBI backbone (NCBIP:139777)
C;Genetics:
A;Gene: GDB:MPZ; CMT1B; CMT1
A;Cross-references: GDB:125266; OMIM:159440
A;Map position: 1q22-1q23
C;Superfamily: myelin P0 protein; immunoglobulin homology
F:43-129/Domain: immunoglobulin homology <IMM>

Query Match 14.5%; Score 163.5; DB 2; Length 251;
Best Local Similarity 28.3%; Pred. No. 1.3e-06;
Matches 54; Conservative 33; Mismatches 81; Indels 23; Gaps 7;

QY 11 ASLVLYVSVCFPCVVEPSTEAVQGNPKMLRCISCKREVEATTVVEWFFYRPEGK 70
DB 20 SSLVL-----SPAQAIVVYTDREHVGAVGSRVTLHC-SFWSSEWSDDISFTWRYQPEGGR 74

QY 71 DFL-IYEVNNGHQEVE--SPFGRLQWNGSKDLQDVSTVLNVLNDSGLYTCNVSRFFE 127
DB 75 DAISIFHYAKGPYIDEVGTGFKERIQWVGDPKWDGSIHNLDSYDNGTFTCDVKNPPD 134

QY 128 FEHRPFVKTTLRLPLRVTEAGEDFTSVSVSEIMMYILLVFLTLWLLIEMI-YCY----- 181
DB 135 -----IVGKTSQVTLVYFEKVPTRYGVVIGVGLGVLLVLLVLLVYVRYCWLRRQA 188

QY 182 ---RKVSKAE 189
DB 189 ALQRRLSAMEK 199

RESULT 7
A29128
myelin P0 protein - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C;Accession: A29128
R;Sakamoto, Y.; Kitamura, K.; Yoshimura, T.; Uyemura, K.
J. Biol. Chem. 262, 4208-4214, 1987
A;Title: Complete amino acid sequence of P0 protein in bovine peripheral nerve myelin.
A;Reference number: A29128; MUID:87166035; PMID:2435734
A;Accession: A29128
A;Molecule type: protein
A;Residues: 1-219 <SAK>
C;Superfamily: myelin P0 protein; immunoglobulin homology
C;Keywords: glycoprotein; membrane protein; myelin; Schwann cell; structural protein
```

F:14-100/Domain: immunoglobulin homology <IMM>

Query Match 14.1%; Score 159; DB 1; Length 219;  
 Best Local Similarity 26.5%; Pred. No. 2.8e-06;  
 Matches 52; Conservative 31; Mismatches 99; Indels 14; Gaps 6;  
 QY 25 VCVEPSETEAVQGNPKLRCISCMKREBEATTVEFYRPEGKDFL-IYEYRNHGE 83  
 DB 1 IVVYTDKEVHGVAGSQVTLVC-SFWSSEWSDLSFTWYQPEGRDAISIFHYAKQPY 59  
 QY 84 VE--SPFQRLQWNGSKDLQVSVITVLNDSGLYTCNVSREFEFAHRPFVKTTLRI 141  
 DB 60 IDEVGTFRQIQWGDGPHKQSGIVIHNLVDYDNGTFTCDVKNPPD-----IVGKTSQV 113  
 QY 142 PLRVTEEAGEDTSVVSSEIMVILLVFTLMLLIEMI-YCYRVKSKAEFA--AQENASD 197  
 DB 114 TLYVEKVPTRYGVVLGAVIGVGLVGLVALLLFLYRYCYRLRQAQLRLHAMEGKL 173  
 QY 198 YLAIPSENKENSAPV 213  
 DB 174 HKTAKDASKRGQTPV 199

# RESULT 8

A54662  
 myelin p0 protein - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jan-2000  
 C:Accession: A54662  
 R:You, K.H.; Hsieh, C.L.; Hayes, C.; Stahl, N.; Francke, U.; Popko, B.  
 Genomics 9, 751-757, 1991  
 A:Title: DNA sequence, genomic organization, and chromosomal localization of the mouse p0  
 A:Reference number: A54662; MUID:91244320; PMID:1709914  
 A:Accession: A54662  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-247 <YOU>  
 A:Cross-references: GB:M62427  
 C:Superfamily: myelin p0 protein; immunoglobulin homology  
 F:43-128/Domain: immunoglobulin homology <IMM>

Query Match 13.7%; Score 154.5; DB 1; Length 247;  
 Best Local Similarity 27.6%; Pred. No. 7.6e-06;  
 Matches 58; Conservative 34; Mismatches 99; Indels 19; Gaps 8;  
 QY 11 ASVLIVYVSVCFPVCEVPESETEAVQGNPKLRCISCMKREBEATTVEFYRPEGK 70  
 DB 20 SSLVL----SPALIVYVTDRIYGVAGSQVTLHC-SFWSSEWSDLSFTWYQ-EGGR 73  
 QY 71 DFL-IYEYRNHGEVE--SPFQRLQWNGSKDLQVSVITVLNDSGLYTCNVSREFE 127  
 DB 74 DAISIPHYAKQPYIDEVGAFFERIQWGDGPHKQSGIVIHNLVDYDNGTFTCDVKNPPD 133  
 QY 128 FEHRPFVKTTLRIPLRVTEEAGEDTSVVSSEIMVILLVFTLMLLIEMI-YCYRVKSK 186  
 DB 134 -----IVGKTSQVTLVVEKVPTRYGVVLGAVIGVGLVALLLFLYRYCYRLRQA 187  
 QY 187 AEE--AAQENASDYLAIPESENKENSAPV 213  
 DB 188 ALQRLSAMEKGRFKHKSXDSKRGQTPV 217

# RESULT 9

A32999  
 myelin p0 protein precursor - horn shark  
 C:Species: Heterodontus francisci (horn shark)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: A32999  
 R:Saavedra, R.A.; Fors, L.; Abersold, R.H.; Arden, B.; Horvath, S.; Sanders, J.; Hood, J.  
 J. Mol. Evol. 29, 149-156, 1989  
 A:Title: The myelin proteins of the shark brain are similar to the myelin proteins of th  
 A:Reference number: A32999; MUID:90040744; PMID:2478717  
 A:Accession: A32999

A:Status: preliminary  
 A:Molecule type: mRNA

A:Residues: 1-246 <SAA>  
 A:Cross-references: GB:X16714; NID:963976; PIDN:CAB37865.1; PID:94467434  
 C:Superfamily: myelin p0 protein; immunoglobulin homology  
 C:Keywords: glycoprotein; membrane protein; myelin; Schwann cell; structural protein  
 F:41-127/Domain: immunoglobulin homology <IMM>

Query Match 12.9%; Score 144.5; DB 1; Length 246;  
 Best Local Similarity 25.8%; Pred. No. 5.4e-05;  
 Matches 49; Conservative 37; Mismatches 77; Indels 27; Gaps 7;  
 QY 16 IYVSVVCFPVCEVPESETEAVQ-----GNPKLRCISCMKREBEATTVEFYRFP 66  
 DB 10 LFCCSVLYAFSLRPSQGISVSTHNLHKTGSDVLYC-GFWSNEYVSDLTLSWRFP 68  
 QY 67 EGKDFL-IYEYRNHGEVE--SPFQRLQWNGSKDLQVSVITVLNDSGLYTCNVS 123  
 DB 69 DNSRDIISIFHYGVNGVPIEKWQFGRVWGDLSKDGSIIVIRNLDYDNGTFTCDVK 128  
 QY 124 REFEFAHRPFVKTTLRIPLRVTEE-----AGEDTSVVSSEIMVILLVFTLMLLIEMI 178  
 DB 129 NPPD-----VVGTSVDVHLTYDKIPPVGVAGVWSGAILGFIILLVGGVLYL--FR 179  
 QY 179 YCYRVKSKAE 188  
 DB 180 YIVRRARSE 189

# RESULT 10

A61087  
 myelin p0 glycoprotein precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jan-2000  
 C:Accession: A61087  
 R:Barbu, M.  
 J. Neurosci. Res. 25, 143-151, 1990  
 A:Title: Molecular cloning of cDNAs that encode the chicken p0 protein: evidence for e  
 A:Reference number: A61087; MUID:90204597; PMID:1690817  
 A:Accession: A61087  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-249 <BAR>  
 C:Comment: This protein is found only in peripheral nervous system Schwann cells.  
 C:Superfamily: myelin p0 protein; immunoglobulin homology  
 C:Keywords: glycoprotein; myelin; Schwann cell; structural protein; transmembrane prote  
 F:1-29/Domain: signal sequence #status predicted <SIG>  
 F:30-249/Product: myelin p0 glycoprotein #status predicted <MAT>  
 F:30-133/Domain: extracellular #status predicted <EXT>  
 F:43-129/Domain: immunoglobulin homology <IMM>  
 F:154-179/Domain: transmembrane #status predicted <TM>  
 F:180-249/Domain: intracellular #status predicted <INT>  
 F:122/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 12.5%; Score 140.5; DB 1; Length 249;  
 Best Local Similarity 23.6%; Pred. No. 0.00012;  
 Matches 48; Conservative 43; Mismatches 87; Indels 25; Gaps 8;  
 QY 13 LVLIYVSVVCFP-----VCVEPSETEAVQGNPKLRCISCMKREBEATTVEFYRPE 67  
 DB 13 LLVLGLLSASGPSPTLAIHVYTPREVYGVTSVHTLSC-SFWSSEWSDLSFTWHPQAE 71  
 QY 68 GKKDFL-IYEYRNHGEVE--SPFQRLQWNGSKDLQVSVITVLNDSGLYTCNVS 124  
 DB 72 GSRDSISIFHYGVNGVPIEKWQFGRVWGDLSKDGSIIVIRNLDYDNGTFTCDVK 131  
 QY 125 EPEFEFAHRPFVKTTLRIPLRVTEEAGEDTSVVSSEIMVILLVFTLMLLIEMI-YCY-- 181  
 DB 132 PPD-----IVKSSQVTLVLEKVPTRYGVVLGAVIGVGLVALLVAVVILVRFCL 185  
 QY 182 -----RVKSKAEFA-AQENASD 197  
 DB 186 RQAVLQRLSAMEKGRKQRLQSAKD 208

## RESULT 11

A57843 sodium channel beta 2 subunit - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: A57843  
R:Issm, L.L.; Ragsdale, D.S.; DeJongh, K.S.; Westebroek, R.E.; Reber, B.F.X.; Scheuer, Cell 83, 433-442, 1995  
A:Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels, a Transmembrane Protein  
A:Reference number: A57843; MUID:96067641; PMID:8521473  
A:Accession: A57843  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-215 <RES>  
A:Cross-references: EMBL:U37026; NID:g1086496; PIDN:AAC52967.1; PID:g1086497  
C:Genetics: SCNB2  
A:Gene: SCNB2

Query Match 10.5%; Score 118; DB 2; Length 215;  
Best Local Similarity 25.1%; Pred. No. 0.0083;  
Matches 49; Conservative 41; Mismatches 85; Indels 20; Gaps 8;  
QY 27 VEVPESETEAVQGNPKLRCL--SCMKREEVATTVWFYRPEG--KDFLIYEVNGH 81  
DB 32 VTVPFTLSVLNGSDTRLPCTNSCYTVNHQFS--LWNTYQCSNCSSEMFQFRMKILN 89  
QY 82 QEVESPFQGRLOWNGSKDLOVSTVLNVTINDSLGTYCNVSRFEFEFAHRPFVKTRLI 141  
DB 90 LKLER-FGDRVFESGNSKYDVSVTLNQVLEDEGIYNCYITN--PPDRHRGHGK---I 142  
QY 142 PLRVTEAGEFTSVSVSEIMMYILLVLTLLIEMIV-CYRKVKAEAEAAQENASDYLA 200  
DB 143 YLQVLLVPPERDSTVAIVGASVGGFLAVVILVMVVKVRR-----EAIYCGVFLIACMV 197  
QY 201 IPSENKENSAPVVE 215  
DB 198 TEEGKTDGEGNAED 212

## RESULT 12

I61783 sodium channel beta 2 subunit - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I61783  
R:Issm, L.L.; Ragsdale, D.S.; DeJongh, K.S.; Westebroek, R.E.; Reber, B.F.X.; Scheuer, Cell 83, 433-442, 1995  
A:Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels, a Transmembrane Protein  
A:Reference number: A57843; MUID:96067641; PMID:8521473  
A:Accession: I61783  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-166 <RES>  
A:Cross-references: EMBL:U37147; NID:g1086498; PIDN:AB60506.1; PID:g1086499  
C:Genetics: SCNB2  
A:Gene: SCNB2  
A:Introns: 50/3; 120/2

Query Match 10.4%; Score 117; DB 2; Length 186;  
Best Local Similarity 24.0%; Pred. No. 0.0087;  
Matches 49; Conservative 38; Mismatches 79; Indels 38; Gaps 8;  
QY 27 VEVPESETEAVQGNPKLRCL--SCMKREEVATTVWFYRPEGKDFLIYEVNGHQEV 84  
DB 3 VTVPFTLSVLNGSDTRLPCTNSCYTVNHQFS--LWNTYQ-----ECSNCSSEEM 50  
QY 85 ESFPQ-----GRLQWNGSKDLOVSTVLNVTINDSLGTYCNVSRFEFEFAHR 132  
DB 51 VLQFRMKINLKLERFGDRVFSGNSKYDVSVTLNQVLEDEGIYNCYITN--PPDRHR 108  
QY 133 PRVKTTRILPLRVTEAGEFTSVSVSEIMMYILLVLTLLIEMIV-CYRKVKAEAEAA 191

DB 109 GHGK---IYQVLLVPPERDSTVAIVGASVGGFLAVVILVMVVKVRR-----KKE 159

QY 192 QENASDYLAIPSENKENSAPVVE 215

DB 160 QKLSTDDLKTEEGKTDGEGNAED 183

## RESULT 13

S17295 fibroblast growth factor receptor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Jul-1999  
C:Accession: S17295  
R:Raz, V.; Kelman, Z.; Avi, A.; Neufeld, G.; Givol, D.; Yarden, Y. Oncogene 6, 753-760, 1991  
A:Title: PCR-based identification of new receptors: molecular cloning of a receptor for fibroblast growth factor  
A:Reference number: S17295; MUID:91270892; PMID:17111190  
A:Accession: S17295  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-820 <RAZ>  
A:Cross-references: EMBL:X55441; NID:G50141; PIDN:CAA39083.1; PID:G50142  
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein kinase domain  
F:478-763/Domain: protein kinase homology <KIN>

Query Match 9.9%; Score 111.5; DB 2; Length 820;  
Best Local Similarity 25.9%; Pred. No. 0.12;  
Matches 41; Conservative 23; Mismatches 55; Indels 39; Gaps 7;

QY 83 EVESPFQGRLOWNGSKDLOVSTVLNVTINDSLGTYCNVSRFEFEFAHRPFVKTTRIL 116

DB 278 KYSDAQPHIQRKHVKGNGSKYDGLFYLKVLKAGVNTDKSEIVLIRNVTPEAG 337

QY 117 LYTCNVSRFEFEFAHRPFVKTTRILPLRVTE-EAGEDFTSVSEIMMYILLVLTLLI 175

DB 338 EYCLAGNSIGISFHSAWL-TVLPAPVREKEITAPDYL-----EAIYCGVFLIACMV 392

QY 176 EMIVCYRKVKAEAEAAQENASDYLAIPSENKENSAPV 213

DB 393 TVIFCRMTTKK-----PDFSSQPAVHKLTKRIFL 423

## RESULT 14

TWMSBK fibroblast growth factor receptor precursor - mouse  
N:Alternate names: bek transforming protein; fibroblast growth factor receptor 2; keratin 10  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1991 #sequence\_revision 13-Mar-1997 #text\_change 16-Jul-1999  
C:Accession: A44142; A31378  
R:Manukhian, A.; Dell'Era, P.; Moscatelli, D.; Kornbluth, S.; Hanafusa, H.; Basilico, Proc. Natl. Acad. Sci. U.S.A. 89, 3305-3309, 1992  
A:Title: Characterization of the murine BEK fibroblast growth factor (FGF) receptor: a novel tyrosine kinase identified by phosphotyrosine antibody screening of cDNA libraries  
A:Reference number: A44142; MUID:92228773; PMID:11373495  
A:Accession: A44142  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-821 <MAN>  
A:Cross-references: GB:M86441  
R:Kornbluth, S.; Paulson, K.E.; Hanafusa, H. Mol. Cell. Biol. 8, 5541-5544, 1988  
A:Title: Novel tyrosine kinase identified by phosphotyrosine antibody screening of cDNA libraries  
A:Reference number: A31378; MUID:89219016; PMID:2468999  
A:Accession: A31378  
A:Molecule type: mRNA  
A:Residues: 477-821 <KOR>  
A:Cross-references: GB:M23362; NID:G533219; PIDN:AAA37285.1; PID:G533220  
C:Genetics: SCNB2  
A:Gene: bek  
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein kinase domain  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor

F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-23/Product: fibroblast growth factor receptor bek #status predicted <MAT>  
F:172-233/Domain: immunoglobulin homology <IM>  
F:378-398/Domain: transmembrane #status predicted <TM>  
F:479-764/Domain: protein kinase homology <KIN>  
F:487-495/Region: protein kinase ATP-binding motif  
F:62-107,179-231,278-342/Disulfide bonds: #status predicted  
F:83,123,147,241,265,297,318,331/Binding site: carboxylate (Asn) (covalent) #status pre  
F:517,534,626/Active site: Lys, Glu, Asp #status predicted  
F:657/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 9.8%; Score 111.5; DB 1; Length 821;  
Best/Local Similarity 25.9%; Pred. No. 0.12;  
Matches 41; Conservative 23; Mismatches 55; Indels 39; Gaps 7;

QY 83 EYSPFQGLQW-----NGSK-----DLQDVSTVL---NVLNDSG 116  
Db 279 KYSDAQPHIQWIKHVEKNGSKYGDGLPVLKVLKAGVNTTDEIEVLIRNVTEDAG 338  
QY 117 LYTCNVSRFEFEAHRFPVKTRILPLRVTE-EGEDFTSVSEIMMYILLVLTLLLI 175  
Db 339 EYTCLAGNSIGISFHSWL-TVLPAPVREKEITASPDYL---EIAIYCIQGVFLIACMVV 393  
QY 176 EMICYKVKVKAERAAQENASDYLAIPESENKENSAPV 213  
Db 394 TVIFRMKTTTK-----PDFSSQPAVHKLTKRIPL 424

## RESULT 15

B36182  
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type PTP69D precursor - fruit fly  
N/A:Alternate names: protein-tyrosine-phosphatase DPTP  
C:Species: Drosophila melanogaster  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: B36182  
R:Streuli, M.; Krueger, N.X.; Tsai, A.Y.M.; Saito, H.  
Proc. Natl. Acad. Sci. U.S.A. 86, 8698-8702, 1989  
A:Title: A family of receptor-linked protein tyrosine phosphatases in humans and Drosoph  
A:Reference number: A36182; MUID:90046860; PMID:2554325  
A:Accession: B36182  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1462 <STR>  
A:Cross-references: GB:M27699; NID:g158188; PIDN:AAA28842.1; PID:g158189  
C:Genetics:  
A:Gene: Ptp69D  
A:Cross-references: FlyBase:FBgn0014007  
C:Superfamily: protein-tyrosine-phosphatase, receptor type PTP69D; fibronectin type III  
phosphatase homology  
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-1462/Product: protein-tyrosine-phosphatase, receptor type PTP69D #status predicted  
F:29-806/Domain: extracellular #status predicted <EXT>  
F:38-114/Domain: immunoglobulin homology <IM>  
F:147-216/Domain: immunoglobulin homology <IM2>  
F:807-823/Domain: transmembrane #status predicted <TM>  
F:824-1462/Domain: intracellular #status predicted <INT>  
F:917-1145/Domain: protein-tyrosine-phosphatase homology <PTP1>  
F:1213-1439/Domain: protein-tyrosine-phosphatase homology <PTP2>  
F:45-112,154-214/Disulfide bonds: #status predicted  
F:1097/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1103/Binding site: substrate phosphate (Arg) #status predicted  
F:1391/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1397/Binding site: substrate phosphate (Arg) #status predicted

Query Match 9.8%; Score 110.5; DB 1; Length 1462;  
Best/Local Similarity 26.7%; Pred. No. 0.28;  
Matches 47; Conservative 31; Mismatches 61; Indels 37; Gaps 11;

QY 4 FNRLEPLASLVLYNWSVCFVCPVSETE-AVQGNPMKLRICISCMKREVEATTVEW 62  
Db 5 YRMSMLNIIILAY-IFLC-AICVQGSVKQEWASIGKNVSLCAS-----ENEA---VAW 54

QY 63 FYRPEG-GKDFLIYEYRNHGHQEVESPPQGRQLQWNGSKDLQDVS-----ITVLNVLNDSG 116  
Db 55 KLGNOTINKNHTRYKIRT-----EPLKSNDDGSENNDSDFMKYKNVLTLDDVINDSG 108  
QY 117 LYTC-----NVSREFEFEAHRFPVKTRILPLRVTEEGEDFTSVSEIMMYILL 166  
Db 109 NYTCTAQGGQNHSTEFQVKFPLP-SKVLQSTPDRIKKIKOD-----VMDYCLI 156

Search completed: December 5, 2003, 23:23:26  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 20:51:34 / Search time 26 Seconds  
(without alignments)  
388.875 Million cell updates/sec

Title: US-09-936-680-2

Perfect score: 1124  
Sequence: 1 MPANRLFLPLASLVLYWVS.....SDYLAIPSENKENSAPVVEE 215

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1124	100.0	1 CIB3 HUMAN	Q9NY72 homo sapien
2	477	42.4	1 CIB1 RABIT	P53788 corytolagus
3	472	42.0	1 CIB1 RAT	Q00954 rattus norv
4	471	41.9	1 CIB1 HUMAN	Q07699 homo sapien
5	468	41.6	1 CIB1 MOUSE	P97952 mus musculus
6	175	15.6	1 MYPO RAT	P06907 rattus norv
7	172	15.3	1 MYPO MOUSE	P25173 mus musculus
8	163.5	14.5	1 MYPO HUMAN	P25189 homo sapien
9	159	14.1	1 MYPO BOVIN	P10522 bos taurus
10	149	13.3	1 EVAL MOUSE	Q70255 mus musculus
11	147	13.1	1 EVAL HUMAN	O60487 homo sapien
12	144.5	12.9	1 MYPO HETFR	P20938 heterodontu
13	140.5	12.5	1 MYPO CHICK	P37301 gallus gall
14	118	10.5	1 CIB2 RAT	P54900 rattus norv
15	111.5	9.9	1 RGR2 MOUSE	P21803 mus musculus
16	110.5	9.8	1 CIB2 HUMAN	O60939 homo sapien
17	110.5	9.8	1 FRP6 DROME	P16620 drosophila
18	107	9.5	1 FRP3 MOUSE	O61851 mus musculus
19	106	9.4	1 TR18 MOUSE	O61098 mus musculus
20	101.5	9.0	1 CEK2 CHICK	P18460 gallus gall
21	99.5	8.9	1 FRG2 HUMAN	P21802 homo sapien
22	99	8.8	1 TR18 HUMAN	Q13478 homo sapien
23	98.5	8.8	1 CEK3 CHICK	P18461 gallus gall
24	94.5	8.4	1 A33 HUMAN	Q99795 homo sapien
25	93.5	8.3	1 LV6B HUMAN	P01722 homo sapien
26	93.5	8.3	1 TVA1 HUMAN	P04436 homo sapien
27	93	8.3	1 FGR3 HUMAN	P22607 homo sapien
28	92	8.2	1 PAS2 SCHAM	P22648 schistocerc
29	90	8.0	1 CXAR HUMAN	P78310 homo sapien
30	89	7.9	1 FGR2 XENLA	Q03364 xenopus lae
31	88.5	7.9	1 CXAR MOUSE	P97792 mus musculus
32	87.5	7.8	1 C07 MOUSE	P50283 mus musculus
33	86.5	7.7	1 UL01 HCMVA	P16719 human cytom

34 86.5 7.7 298 1 JAM1 BOVIN  
35 85.5 7.6 503 1 SHS1 HUMAN  
36 85 7.6 273 1 BAS1 MOUSE  
37 85 7.6 605 1 RSK2 GEOCY  
38 84.5 7.5 261 1 HB23 HUMAN  
39 84 7.5 108 1 KV50 MOUSE  
40 84 7.5 205 1 CD83 HUMAN  
41 83.5 7.4 133 1 HV2F HUMAN  
42 83.5 7.4 268 1 KB2X HUMAN  
43 83.5 7.4 654 1 BFR2 HUMAN  
44 83 7.4 417 1 PVR HUMAN  
45 83 7.4 490 1 YVO1 CAEEL

## ALIGNMENTS

### RESULT 1

CIB3 HUMAN  
ID CIB3 HUMAN STANDARD; PRT; 215 AA.  
AC Q9NY72;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Sodium channel beta-3 subunit precursor.  
GN SCN3B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=20160948; PubMed=10688874; Cox P., Dixon A.K., Lee K., Morgan K., Stevens E.B., Shaw B., Richardson P.J., Mizuguchi K., Jackson A.P.; Pinnock R.D., Hughes J., Richardson P.J., Mizuguchi K., Jackson A.P.; "Beta3: an additional auxiliary subunit of the voltage-sensitive sodium channel that modulates channel gating with distinct kinetics."; Proc. Natl. Acad. Sci. U.S.A. 97:2308-2313(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Amygdala;  
RX MEDLINE=21154917; PubMed=11230166; Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S., Ansgore W., Boecker M., Bloecker H., Bauersachs S., Blum H., Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N., Weses H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D., Wambutt R., Korn B., Klein M., Foustka A.; "Towards a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs."; Genome Res. 11:422-435(2001).  
CC -!- FUNCTION: MODULATES CHANNEL GATING KINETICS.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -!- EMBL; AJ243396; CAB76825.1; -!  
CC HSSP; P06907; 1NEU.  
CC Genew; HGNC:20665; SCN3B.  
CC GO; GO:0016021; C:integral to membrane; NAS.  
CC GO; GO:0005248; F:voltage-gated sodium channel activity; NAS.  
CC GO; GO:0006814; F:sodium ion transport; NAS.  
CC InterPro; IPR007110; IG-like.  
CC InterPro; IPR003599; IG.  
CC InterPro; IPR003006; IG\_MHC.

DR PFAM: PF00047; ig: 1.  
 DR SMART: SM00409; ig: 1.  
 DR PROSITE: PS0635; IG\_LIKE; 1.  
 KW Ionic channel; transmembrane; Ion transport; Voltage-gated channel;  
 KW Glycoprotein; Signal; Immunoglobulin domain.  
 FT SIGNAL 1 22 BY SIMILARITY.  
 FT CHAIN 23 215 SODIUM CHANNEL BETA-3 SUBUNIT.  
 FT DOMAIN 23 215 SODIUM CHANNEL BETA-3 SUBUNIT.  
 FT TRANSMEM 160 180 POTENTIAL.  
 FT DOMAIN 181 215 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 32 154 IG-LIKE C2-TYPE.  
 FT DISULFID 45 120 POTENTIAL.  
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 215 AA; 24702 MW; 70F6C604E9E26662 CRC64;

Query Match 100.0%; Score 1124; DB 1; Length 215;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-94;  
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPANRLPLASLVLYVSVCFPCVPEVPESEAVQGNPKLRCISCMKREVEATTVV 60  
 DB 1 MPANRLPLASLVLYVSVCFPCVPEVPESEAVQGNPKLRCISCMKREVEATTVV 60

QY 61 EWFYRPEGKDFLIYVRNGHQEVESPQGRLOWNGSKLDQVSVITVNLNDGLTYC 120  
 DB 61 EWFYRPEGKDFLIYVRNGHQEVESPQGRLOWNGSKLDQVSVITVNLNDGLTYC 120

QY 121 NVSRFEFEARHPFKTKTRILPLRVTEAGDEFTSVSEIMMYLLVFLTLWLLIEMIYC 180  
 DB 121 NVSRFEFEARHPFKTKTRILPLRVTEAGDEFTSVSEIMMYLLVFLTLWLLIEMIYC 180

QY 181 YRKSKAEAAQENASDYLAIPESENKENSAPVVEE 215  
 DB 181 YRKSKAEAAQENASDYLAIPESENKENSAPVVEE 215

RESULT 2  
 ID CIB1 RABIT STANDARD; PRT; 218 AA.  
 AC P53788;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Sodium channel beta-1 subunit precursor.  
 GN SCN1B.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1] "Cloning of the cDNA encoding the sodium channel beta 1 subunit from rabbit."  
 RC STRAIN=New Zealand white; TISSUE=Sciatic nerve;  
 RX MEDLINE=96235151; PubMed=8666261;  
 RA Belcher S.M., Howe J.R.;  
 RT "Cloning of the cDNA encoding the sodium channel beta 1 subunit from rabbit."  
 RL Gene 170:285-286 (1996).  
 CC -!- FUNCTION: CRUCIAL IN THE ASSEMBLY, EXPRESSION AND FUNCTIONAL MODULATION OF THE HETEROTRIMERIC COMPLEX OF THE SODIUM CHANNEL. THE BETA-1 SUBUNIT CAN MODULATE MULTIPLE ALPHA SUBUNIT ISOFORMS FROM BRAIN, SKELETAL MUSCLE, AND HEART.  
 CC -!- SUBUNIT: THE SODIUM CHANNEL CONSISTS OF A PORE-FORMING ALPHA SUBUNIT, BETA-1 AND BETA-2 SUBUNITS. BETA-1 IS NONCOVALENTLY ASSOCIATED WITH ALPHA, WHILE BETA-2 IS COVALENTLY LINKED BY DISULFIDE BONDS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U35382; AAB17572.1; -.  
 CC PIR: J04788; J04788.  
 CC InterPro: IPR003006; IG\_MHC.  
 CC Pfam: PF00047; ig: 1.  
 CC PROSITE: PS0635; IG\_LIKE; FALSE\_NEG.  
 KW Ionic channel; transmembrane; Ion transport; Voltage-gated channel;  
 KW Glycoprotein; Signal; Immunoglobulin domain.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 218 SODIUM CHANNEL BETA-1 SUBUNIT.  
 FT DOMAIN 19 160 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 161 182 POTENTIAL.  
 FT DOMAIN 183 218 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 22 150 IG-LIKE C2-TYPE.  
 FT DISULFID 40 121 POTENTIAL.  
 FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 218 AA; 24706 MW; 39BD174E1FAE7FD2 CRC64;

Query Match 42.4%; Score 477; DB 1; Length 218;  
 Best Local Similarity 49.5%; Pred. No. 9.3e-36;  
 Matches 106; Conservative 32; Mismatches 68; Indels 8; Gaps 5;

QY 10 LASLVLYVSVCFPCVPEVPESEAVQGNPKLRCISCMKREVEATTVVFPYRPEG 69  
 DB 5 LAFVGAALVSAAGCGVEVDSEAVGTMFKILCSCKRSETTAEFTETFRQXGT 64

QY 70 KDFL-ITYRNGHQEVESP--FQGRLOWNGS---KDLQDVSVITVNLNDGLTYCNVS 123  
 DB 65 BEFVKILYENEVQLQEDDERFEGRVVNGSRGKDLQDLISIPITVNTVYHNSGDYQCHVY 124

QY 124 REFEEARHPFKTKTRILPLRVTEAGDEFTSVSEIMMYLLVFLTLWLLIEMIYC 183  
 DB 125 RLSEFENYHTSVVKILHLEVDKANDMASIVSEIMMYLLVFLTLWLLIEMIYC 184

QY 184 VSKA-EAAQENASDYLAIPESENKENSAPVVEE 215  
 DB 185 IAAATEAAQENASEYLAIITSEKNCCTGVQVAE 218

RESULT 3  
 ID CIB1 RAT STANDARD; PRT; 218 AA.  
 AC Q00954;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Sodium channel beta-1 subunit precursor.  
 GN SCN1B.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1] "Cloning of the cDNA encoding the sodium channel beta 1 subunit of the rat brain sodium channel."  
 RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=92271207; PubMed=1375395;  
 RA Isom L.L., De Jongh K.S., Patton D.E., Reber B.F.X., Offord J., Charbonneau H., Walsh K., Goldin A.L., Catterall W.A.;  
 RT "Primary structure and functional expression of the beta 1 subunit of the rat brain sodium channel."  
 RL Science 256:839-842 (1992).  
 CC -!- FUNCTION: CRUCIAL IN THE ASSEMBLY, EXPRESSION, AND FUNCTIONAL MODULATION OF THE HETEROTRIMERIC COMPLEX OF THE SODIUM CHANNEL. THE BETA-1 SUBUNIT CAN MODULATE MULTIPLE ALPHA SUBUNIT ISOFORMS FROM BRAIN, SKELETAL MUSCLE, AND HEART.

CC -!- SUBUNIT: THE SODIUM CHANNEL CONSISTS OF A PORE-FORMING ALPHA  
CC SUBUNIT, BETA-1 AND BETA-2 SUBUNITS. BETA-1 IS NONCOVALENTLY  
CC ASSOCIATED WITH ALPHA, WHILE BETA-2 IS COVALENTLY LINKED BY  
CC DISULFIDE BONDS.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART, SKELETAL MUSCLE,  
CC AND SPINAL CORD.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M91808; AAA88513.1; -;  
CC PIR; A42737; A42737;  
CC InterPro; IPR003006; Ig\_MHC.  
CC Pfam; PF00047; Ig\_1.  
CC PROSITE; PS00835; IG\_LIKE; FALSE NEG.  
CC Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;  
CC Glycoprotein; Signal; Immunoglobulin domain.  
CC SIGNAL 1 18  
CC FT CHAIN 19 218  
CC FT DOMAIN 19 160  
CC FT TRANSMEM 161 182  
CC FT DOMAIN 183 218  
CC FT DOMAIN 22 150  
CC FT DISULFID 40 121  
CC FT CARBOHYD 93 93  
CC FT CARBOHYD 110 110  
CC FT CARBOHYD 114 114  
CC FT CARBOHYD 135 135  
CC SEQUENCE 218 AA; 24692 MW; 08A84FC44FF23068 CRC64;  
CC  
CC Query Match  
CC Best Local Similarity 49.1%; Pred. No. 2.6e-35;  
CC Matches 105; Conservative 33; Mismatches 68; Indels 8; Gaps 5;  
CC  
CC QY 10 LASLVLTWVSVCFFVCEVSETEAVQCNPKLCISCKRREVEATTVVEFYRPEGG 69  
CC DB 5 LALVVGAVLVSWGCGVEVDSETEAVYGMTFKLCISCKRSETTAEFTTFKRGKT 64  
CC  
CC QY 70 KDFL-ITYRNHGHQVESP--FQGRLOWNGS---KQLQDVSIITLVNTLNDSGLYTCNVS 123  
CC DB 65 EEFVKILAYENEVLQLEDERFEGRVVWNGSRGKQLQLDLSIFITVTVYHSGDYECVY 124  
CC  
CC QY 124 REFTEAHRPFVKTTRILPLAVTEBAGDFTSVSVSEINMYLLVFLTLWLLIEMICYRK 183  
CC DB 125 RLLFFDFNYHTSVVVKTHLEVDVKANRDMASIVSEINMYLLVFLTLWLLIEMICYRK 184  
CC  
CC QY 184 VSKA-EEAQAENASDYLAIPESENKEN-SAVPEE 215  
CC DB 185 TAAATEAAQAENASDYLAIPESENKENCTGVQVAE 218  
CC  
CC RESULT 4  
CC CIB1 HUMAN  
CC ID CIB1 HUMAN STANDARD; PRT; 218 AA.  
CC AC Q07659.  
CC DT 01-OCT-1996 (Rel. 34, Created)  
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
CC DE Sodium channel beta-1 subunit precursor.  
CC GN SCN1B.  
CC OS Homo sapiens (Human).  
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
CC OX NCBI\_TaxID=9606;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.

RC TISSUE=Brain;  
RX MEDLINE=93357746; PubMed=8394762;  
RA McClatchey A.I.; Cannon S.C.; Slaughterhaupt S.A.; Guseella J.F.;  
RT "The cloning and expression of a sodium channel beta 1-subunit cDNA  
RT from human brain.";  
RL Hum. Mol. Genet. 2:745-749(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Heart, and Skeletal muscle;  
RX MEDLINE=94171787; PubMed=8125980;  
RA Makita N., Bennett P.B. Jr.; George A.L. Jr.;  
RT "Voltage-gated Na+ channel beta 1 subunit mRNA expressed in adult  
RT human skeletal muscle, heart, and brain is encoded by a single  
RT gene.";  
RL J. Biol. Chem. 269:7571-7578(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=95154833; PubMed=7851891;  
RA Makita N., Sloan-Brown K., Weghuis D.O., Ropers H.-H.,  
RA George A.L. Jr.;  
RT "Genomic organization and chromosomal assignment of the human  
RT voltage-gated Na+ channel beta 1 subunit gene (SCN1B).";  
RL Genomics 23:628-634(1994).  
RN [4]  
RP VARIANT GEFS+ TRP-121.  
RX MEDLINE=98361163; PubMed=9697698;  
RA Wallace R.H., Wang D.W., Singh R., Scheffer I.B., George A.L. Jr.,  
RA Phillips H.A., Saar K., Reis A., Johnson E.W., Sutherland G.R.,  
RA Berkovic S.F., Mulley J.C.;  
RT "Febrile seizures and generalized epilepsy associated with a mutation  
RT in the Na(+)-channel beta-1 subunit gene SCN1B.";  
RL Nat. Genet. 19:366-370(1998).  
CC -!- FUNCTION: CRUCIAL IN THE ASSEMBLY, EXPRESSION, AND FUNCTIONAL  
CC MODULATION OF THE HETEROTRIMERIC COMPLEX OF THE SODIUM CHANNEL.  
CC THE BETA-1 SUBUNIT CAN MODULATE MULTIPLE ALPHA SUBUNIT  
CC ISOFORMS FROM BRAIN, SKELETAL MUSCLE, AND HEART.  
CC -!- SUBUNIT: THE SODIUM CHANNEL CONSISTS OF A PORE-FORMING ALPHA  
CC SUBUNIT, BETA-1 AND BETA-2 SUBUNITS. BETA-1 IS NONCOVALENTLY  
CC ASSOCIATED WITH ALPHA, WHILE BETA-2 IS COVALENTLY LINKED BY  
CC DISULFIDE BONDS.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN SKELETAL MUSCLE, HEART  
CC AND BRAIN.  
CC -!- DISEASE: Defects in SCN1B are a cause of generalized epilepsy with  
CC febrile seizures plus (GEFS+). GEFS+ is a disease  
CC characterized by a highly variable phenotype combining febrile  
CC seizures, generalized seizures often precipitated by fever at age  
CC 6 years or more, and partial seizures, with a variable degree of  
CC severity.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L10338; AAA60391.1; -;  
CC EMBL; L16242; AAA61277.1; -;  
CC EMBL; U12193; AAB97608.1; -;  
CC EMBL; U12189; AAB97608.1; JOINED.  
CC EMBL; U12190; AAB97608.1; JOINED.  
CC EMBL; U12191; AAB97608.1; JOINED.  
CC EMBL; U12192; AAB97608.1; JOINED.  
CC PIR; A55734; A55734.  
CC GenBank; HGNC:10586; SCN1B.  
CC MIM; 600235; -;  
CC MIM; 604233; -;  
CC GO; GO:0005248; P:voltage-gated sodium channel activity; TAS.  
CC GO; GO:0006814; P:sodium ion transport; TAS.



```

RP SEQUENCE FROM N.A.
RX MEDLINE=85124601; PubMed=2578885;
RA Lemke G.; Axel R.;
RT "Isolation and sequence of a cDNA encoding the major structural
RL protein of peripheral myelin.";
RN Cell 40:501-508(1985).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90166482; PubMed=2483091;
RA Lemke G.; Lamar E.; Patterson J.;
RT "Isolation and analysis of the gene encoding peripheral myelin
RL protein zero.";
RN Neuron 1:73-83(1988).
[3]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 30-148.
RX MEDLINE=96413553; PubMed=8816707;
RA Shapiro L.; Doyle J.P.; Hensley P.; Colman D.R.; Hendrickson W.A.;
RT "Crystal structure of the extracellular domain from P0, the major
RL structural protein of peripheral nerve myelin.";
RN Neuron 17:435-449(1996).
CC -1- FUNCTION: CREATION OF AN EXTRACELLULAR MEMBRANE FACE WHICH GUIDES
CC THE WRAPPING PROCESS AND ULTIMATELY COMPACTS ADJACENT LAMELLAE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND ONLY IN PERIPHERAL NERVOUS SYSTEM
CC SCHWANN CELLS.
CC -1- PTM: N-LINKED GLYCAN IS SULFATED.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: BELONGS TO THE MYELIN P0 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K03242; AAA41576.1; -.
CC PIR; JQ0622; MPRT0.
CC PDB; 1NEU; 15-MAY-97.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003596; Ig_V.
CC InterPro; IPR000920; Myelin_P0.
CC Pfam; PF00047; Ig; 1.
CC PRINTS; PR00213; MYELINP0.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS00835; IG LIKE; 1.
CC PROSITE; PS00568; MYELIN P0; 1.
CC Myelin; Structural protein; Glycoprotein; Transmembrane;
CC Phosphorylation; Immunoglobulin domain; Signal; 3D-structure.
CC SIGNAL 1 29
CC CHAIN 30 248 MYELIN P0 PROTEIN.
CC DOMAIN 30 153 EXTRACELLULAR.
CC TRANSMEM 154 179
CC DOMAIN 180 248
CC DOMAIN 30 143 CYTOPLASMIC.
CC DISULFID 50 127 IG-LIKE V-TYPE.
CC CARBOHYD 122 122 N-LINKED (GLCNAC... ) (COMPLEX).
CC CONFLICT 43 43 G -> R (IN REF. 1).
CC STRAND 31 33
CC STRAND 37 41
CC TURN 42 43
CC STRAND 46 48
CC STRAND 51 53
CC STRAND 63 70
CC TURN 71 72
CC STRAND 77 83
CC TURN 84 85
CC STRAND 86 89
CC TURN 94 97
CC STRAND 99 101
CC STRAND 104 104

```

```

FT HELIX 105 107
FT TURN 108 108
FT STRAND 109 109
FT STRAND 114 114
FT HELIX 119 121
FT STRAND 123 130
FT STRAND 138 147
SQ SEQUENCE 248 AA; 27570 MW; BB703F173466119B CRC64;
Query Match 15.6%; Score 175; DB 1; Length 248;
Best Local Similarity 28.1%; Pred. No. 1.3e-08;
Matches 59; Conservative 34; Mismatches 99; Indels 18; Gaps 7;
QY 11 ASMLVLYWVSCFVVCVEVSETEAVQGNPMKLCISCIMKREVEATVTVVFRPQGGK 70
DB 20 SSLVL-----SPTLAIVVYTDREYVGVGQSVTLHC-SFWSSEWYSDDISFTWRYQPEGGR 74
QY 71 DFL-IYEVRNGHOEVE--SPFGRLQWNGSKLDQVSIITVLTNDSGLYTCNVSRFE 127
DB 75 DAISIFHYAKQPYIDEVGTGFERIOWGDPGSKGSIIVHNLDSYDSNGTFTCDVKNPDP 134
QY 128 FEHRPFVKTTLPLRVTBAGSDFTSVSEIMMYILLVPLTLMLLIEM-ICYRKVSK 186
DB 135 -----IVGKTSQVTLVYFVKVPTRYGVVLGAVIGGILGVVLLLLLFILIRYCWLRQA 188
QY 187 ASE-----AAQENASDYLAIPSENKENSAPV 213
DB 189 ALQRLSAWEKGFHKSKSKSKRGRQTPV 218

```

```

RESULT 7
MYPO_MOUSE
ID MYPO_MOUSE STANDARD; PRT; 248 AA.
AC P27573;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myelin P0 protein precursor (Myelin protein zero) (Myelin peripheral
DE protein) (NRP).
GN MPZ OR P0.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91244320; PubMed=1709914;
RA You K.H.; Hsieh C.L.; Hayes C.; Stahl N.; Francke U.; Popko B.;
RT "DNA sequence, genomic organization, and chromosomal localization of
RT the mouse peripheral myelin protein zero gene: identification of
RT polymorphic alleles.";
RL Genomics 9:751-757(1991).
CC -1- FUNCTION: CREATION OF AN EXTRACELLULAR MEMBRANE FACE WHICH GUIDES
CC THE WRAPPING PROCESS AND ULTIMATELY COMPACTS ADJACENT LAMELLAE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND ONLY IN PERIPHERAL NERVOUS SYSTEM
CC SCHWANN CELLS.
CC -1- PTM: N-LINKED GLYCAN IS SULFATED.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: BELONGS TO THE MYELIN P0 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M62860; AAA39867.1; -.
CC EMBL; M62857; AAA39867.1; JOINED.
CC EMBL; M62858; AAA39867.1; JOINED.
CC EMBL; M62859; AAA39867.1; JOINED.

```

DR HESP: P06907; 1NEU.  
 DR MGD; MGI:103177; MPZ.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003596; Ig V.  
 DR InterPro: IPR009020; Myelin\_P0.  
 DR PRINTS; PR00213; MYELINP0.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG LIKE; 1.  
 DR PROSITE; PS00568; MYELIN P0; 1.  
 KW Myelin; Structural protein; Glycoprotein; Transmembrane;  
 KW Phosphorylation; Immunoglobulin domain; Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 248 MYELIN P0 PROTEIN.  
 FT DOMAIN 30 153 EXTRACELLULAR (BY SIMILARITY).  
 FT TRANSMEM 154 179 BY SIMILARITY.  
 FT DOMAIN 180 248 CYTOPLASMIC (BY SIMILARITY).  
 FT DOMAIN 30 143 IG-LIKE V-TYPE.  
 FT DISULFID 50 127 POTENTIAL.  
 FT CARBOHYD 122 122 N-LINKED (GLCNAC... ) (COMPLEX).  
 SQ SEQUENCE 248 AA; 27621 MW; 936D66684300CAC9 CRC64;  
 Query Match 15.3%; Score 172; DB 1; Length 248;  
 Best Local Similarity 28.1%; Pred. No. 2 5e-08;  
 Matches 59; Conservative 34; Mismatches 99; Indels 18; Gaps 7;  
 QY 11 ASLVLYWVSVCPCVCEVSETEAVQGNPKLCISCKMREVEATTVVWFVPEGGK 70  
 DB 20 SSILV-----SPALAIIVYTDREIYGAVGSGVTLHC-SFMSSEWVSDDISFTWRYPEGGR 74  
 QY 71 DEL-IVEYRNHGOEVE--SPFGRLQWNGSKDLQDVSIIVLNTLNDGLYTCNVSRFE 127  
 DB 75 DAISIFHYAKQYIDEVGAFRKRIQVGDPRWKGSIIVHNLDSNGTFCVDVKNPFD 134  
 QY 128 FEHRFPVKTPLIRVTEEGEDFTSVSEIMYILVFLMLLIEMI-YCVRYKVK 186  
 DB 135 -----IVGKTSQVTLVPEKVPTRYGVGLGAVIGILGVLLLLFLYIRYCWLRQA 188  
 QY 187 AEE---AAQENASDYLAIPESENKENSAPV 213  
 DB 189 ALQRISAMEKGFHKSDDSKRGQRTV 218  
 RESULT 8  
 MYPO HUMAN STANDARD; PRT; 248 AA.  
 AC P25189; Q16072;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DB Myelin P0 protein precursor (Myelin protein zero) (Myelin peripheral  
 DE protein) (MPP).  
 GN MPZ.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RN RP SEQUENCE FROM N.A.  
 RX MEDLINE=92062068; PubMed=1719967;  
 RA Hayasaka K., Nanao K., Tabara M., Sato W., Takada G., Miura M.,  
 RA Uyemura K.;  
 RT "Isolation and sequence determination of cDNA encoding the major  
 RT structural protein of human peripheral myelin."  
 RL Biochem. Biophys. Res. Commun. 180:515-518(1991).  
 [2]  
 RN RP SEQUENCE FROM N.A., AND VARIANT CMT-1B HIS-98.  
 RC TISSUE=Spinal cord;  
 RX MEDLINE=93356807; PubMed=7688964;  
 RA Hayasaka K., Ohnishi A., Takada G., Fukushima Y., Murai Y.;  
 RT "Mutation of the myelin P0 gene in Charcot-Marie-Tooth neuropathy  
 RT type 1."  
 RL Biochem. Biophys. Res. Commun. 194:1317-1322(1993).  
 [3]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94154677; PubMed=7509228;  
 RA Pham-Dinh D., Fourbil Y., Blanquet F., Mattai M.-G., Roedel N.,  
 RA Latour P., Charot G., Vandenbergh A., Dautigny A.;  
 RT "The major peripheral myelin protein zero gene: structure and  
 RT localization in the cluster of Fc gamma receptor genes on human  
 RT chromosome 1q21.3-q23.";  
 RL Hum. Mol. Genet. 2:2051-2054(1993).  
 [4]  
 RN RP REVIEW ON CMT-1B VARIANTS.  
 RX MEDLINE=95282670; PubMed=7762451;  
 RA Roa B.B., Lupski J.R.;  
 RL "Molecular Genetics of Charcot-Marie-Tooth neuropathy.";  
 RL Adv. Hum. Genet. 22:117-152(1994).  
 [5]  
 RN RP REVIEW ON CMT-1B VARIANTS.  
 RX MEDLINE=94302675; PubMed=7518101;  
 RA Patel P.J., Lupski J.R.;  
 RL "Charcot-Marie-Tooth disease: a new paradigm for the mechanism of  
 RL inherited disease.";  
 RL Trends Genet. 10:128-133(1994).  
 [6]  
 RN RP REVIEW ON CMT-1B AND DSS VARIANTS.  
 RX MEDLINE=99103480; PubMed=9888385;  
 RA Nelis E., Hautes N., van Broeckhoven C.;  
 RT "Mutations in the peripheral myelin genes and associated genes in  
 RT inherited peripheral neuropathies.";  
 RL Hum. Mutat. 13:11-28(1999).  
 [7]  
 RN RP VARIANT CMT-1B MET-30.  
 RX MEDLINE=94061030; PubMed=7694726;  
 RA Hayasaka K., Takada G., Ionasescu V.V.;  
 RT "Mutation of the myelin P0 gene in Charcot-Marie-Tooth neuropathy  
 RT type 1B.";  
 RL Hum. Mol. Genet. 2:1369-1372(1993).  
 [8]  
 RN RP VARIANT CMT-1B CYS-82.  
 RX MEDLINE=94083941; PubMed=7505151;  
 RA Himoto M., Yoshikawa H., Matsui T., Mitsui Y., Takahashi M.,  
 RA Kaido M., Nishimura T., Sawaishi Y., Takada G., Hayasaka K.;  
 RT "New mutation of the myelin P0 gene in a pedigree of  
 RT Charcot-Marie-Tooth neuropathy 1.";  
 RL Biochem. Mol. Biol. Int. 31:169-173(1993).  
 [9]  
 RN RP VARIANTS CMT-1B GLU-90 AND GLU-96.  
 RX MEDLINE=94035113; PubMed=7693129;  
 RA Hayasaka K., Himoro M., Sato W., Takada G., Uyemura K., Shimizu N.,  
 RA Bird T.D., Conneally P.M., Chance P.F.;  
 RT "Charcot-Marie-Tooth neuropathy type 1B is associated with mutations  
 RT of the myelin P0 gene.";  
 RL Nat. Genet. 5:31-34(1993).  
 [10]  
 RN RP VARIANT CMT-1B SER-63 DEL.  
 RX MEDLINE=94035114; PubMed=7693130;  
 RA Kulkens T., Bolhuis P.A., Wolterman R.A., Kemp S., Te Nijenhuis S.,  
 RA Valentijn L.J., Hengels G.W., Jennekens F.G., de Visser M.,  
 RA Hoogendijk J.E., Baas F.;  
 RT "Deletion of the serine 34 codon from the major peripheral myelin  
 RT protein P0 gene in Charcot-Marie-Tooth disease type 1B.";  
 RL Nat. Genet. 5:35-39(1993).  
 [11]  
 RN RP VARIANT CMT-1B GLU-96.  
 RX MEDLINE=94068501; PubMed=7504284;  
 RA Su Y., Brooks D.G., Li L., Lepore J., Trofatter J.A., Ravetch J.V.,  
 RA Lebo R.V.;  
 RT "Myelin protein zero gene mutated in Charcot-Marie-Tooth type 1B  
 RT patients.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10856-10860(1993).  
 [12]  
 RN RP VARIANTS DSS CYS-63 AND ARG-167.  
 RX MEDLINE=94100981; PubMed=7506095;  
 RA Hayasaka K., Himoro M., Sawaishi Y., Nanao K., Takahashi T.,  
 RA Takada G., Nicholson G.A., Ouvrier R.A., Tachi N.;



```

CC -1- TISSUE SPECIFICITY: FOUND ONLY IN PERIPHERAL NERVOUS SYSTEM
CC -1- PTM: N-GLYCOSYLATION IS SULFATED.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MYELIN P0 PROTEIN FAMILY.
DR PIR; A29128; A29128.
DR HSSP; P06907; INEU.
DR GlycoSuiteDB; P10522; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR000920; Myelin_P0.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00213; MYELINP0.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00568; MYELIN_P0; 1.
KW Myelin; structural protein; Glycoprotein; Transmembrane;
KW Phosphorylation; Immunoglobulin domain.
FT DOMAIN 1 124 EXTRACELLULAR.
FT TRANSMEM 125 150
FT DOMAIN 151 219 CYTOPLASMIC.
FT DOMAIN 151 114 IG-LIKE V-TYPE.
FT DISULFID 21 98 POTENTIAL. (GLCNAC. . .) (COMPLEX).
FT CARBOHYD 93 93 N-LINKED. (GLCNAC. . .) (POTENTIAL).
FT MOD_RES 181 181 /FTIC-CAR 000221.
FT MOD_RES 204 204 PHOSPHORYLATION (BY PKC).
FT MOD_RES 214 214 PHOSPHORYLATION (BY PKC).
FT SEQUENCE 219 AA; 24710 MW; E4A882C1B7225FAP CRC64;

Query Match 14.1%; Score 159; DB 1; Length 219;
Best Local Similarity 26.5%; Pred. No. 3.2e-07;
Matches 52; Conservative 31; Mismatches 99; Indels 14; Gaps 6;

QY 25 VCVEVPSSEAVQGNPKLRCISCMKREVEATTVVWFYRPEGKQFL-IYEYRNGHQE 83
DB 1 IVVYTDKEVHGAQSVQVLYC-SFWSSEWVSDLSFTWYQPEGGRDAISIFHYAKGPY 59
QY 84 VE--SPQGLQWNGSKDQVYSITVNLVNDSGLYTCNYSREFEFAHPPFVKTFLI 141
DB 60 IDEVGTFRKIQWGDPHRDRGDSIVHNDYDNGTFCIDVAKNPPD-----IVGKTSQV 113
QY 142 PLRVTEAGSDFTSVVSEIMMVLVFLTLMLIEMI-YCYKRVSKABEA---AGENASD 197
DB 114 TLVFEKVPTRYGVGLGAVGVGVVGLVLLALLFLYLRICMLRRQAQLQRHAMEKGL 173
QY 198 YLAIPSENKENSAPV 213
DB 174 HKTAKASKRGROTPV 189

RESULT 10
ID EVAL_MOUSE STANDARD; PRT; 215 AA.
AC O7025;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Epithelial V-like antigen 1 precursor.
GN EVAL OR EVA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RAG-2; TISSUE=thymus;
RX MEDLINE=98252857; PubMed=9585423;
RA Guttinger M., Sutti F., Panigada M., Porcellini S., Merati B.,
RA Mariani M., Teesalu T., Gonzalez G.G., Grassi F.,
RT "Epithelial V-like antigen (EVA), a novel member of the immunoglobulin
superfamily, expressed in embryonic epithelia with a potential role as

```

```

RT homotypic adhesion molecule in thymus histogenesis.";
RL J. Cell Biol. 141:1061-1071(1998).
CC -1- FUNCTION: Mediates homophilic cell-cell adhesion.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (probable).
CC -1- TISSUE SPECIFICITY: Expressed in liver and gut, skin, and testis
CC but not in thymocytes, lymphocytes, macrophage or dendritic cells
CC or cell lines.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF030454; AAC40128.1; -.
DR HSSP; P06907; INEU.
DR MGD; MGI:1289160; Eva.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR000920; Myelin_P0.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00213; MYELINP0.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Cell adhesion; Immunoglobulin domain; Transmembrane; Glycoprotein;
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 215 EPITHELIAL V-LIKE ANTIGEN 1.
FT DOMAIN 27 154 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 155 175 POTENTIAL.
FT DOMAIN 176 215 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 141 IG-LIKE V-TYPE.
FT DISULFID 47 123 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 215 AA; 24162 MW; F6E5E36787CE69D5 CRC64;

Query Match 13.3%; Score 149; DB 1; Length 215;
Best Local Similarity 23.4%; Pred. No. 2.4e-06;
Matches 47; Conservative 46; Mismatches 100; Indels 8; Gaps 5;

QY 10 LASLVLYVWVYVCPVCEV--PSSTEAVQGNPKLRCISCMKREVEATTVVWFYR- 66
DB 10 LPFLLSQLTALCPTFAVEIYTSGALEAVNGTDLKCTPSSFAPVGDALTVT-WNFRPR 68
QY 67 EGGKD-FLIYEYRNGHQEVESPPQGRLOWNGSKDQVYSITVNLVNDSGLYTCNYSRE 125
DB 69 DGGREQVFVYHMDPPFWSGRFKDRVVDGNPNRYDVSILLKQLQFDNGTTCQVKNP 128
QY 126 FEFBAHRRPFVKTFLIPLRVTEAGSDFTSVVSEIMMVLVFLTLMLIEMIYCYKRV 185
DB 129 PDVGD---LVGTIRLSVHVTFPSEIFPLAVAIGSACALMIIVIVVVLVQHFPRKKNAD 185
QY 186 KAERAAQENASDYLAIPSENK 206
DB 186 RADKAEGTKSKEEKLQNGNK 206

RESULT 11
ID EVAL_HUMAN STANDARD; PRT; 215 AA.
AC O60487;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Epithelial V-like antigen 1 precursor.
GN EVAL OR EVA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX TISSUE=Liver, and Spleen;  
 RX MEDLINE=98252857; PubMed=988423;  
 RA Guttinger M., Sutti F., Panigada M., Porcellini S., Merati B.,  
 RA Marijani M., Teesalu T., Conesalez G.G., Grassi F.,  
 RA "Epithelial V-like antigen (EVA), a novel member of the immunoglobulin  
 RT superfamily, expressed in embryonic epithelia with a potential role as  
 RT homotypic adhesion molecule in thymus histogenesis.";  
 RL J. Cell Biol. 141:1061-1071(1998).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RX TISSUE=Endometrial tumor;  
 RA Davenport J.W., Neale G.A.M., Goorha R.M.;  
 RA "Identification of putative target genes involved in LMO2-induced  
 RT leukemogenesis.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RC SEQUENCE FROM N.A.  
 RX TISSUE=Squamous Cell carcinoma;  
 RA Pietas A., Petersen I., Schluens K., Petersen S.;  
 RA "Human EVL gene which is downregulated in lung carcinoma cells.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RC SEQUENCE FROM N.A.  
 RX TISSUE=Lung;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,  
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Schetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Wuzny K.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Mediates homophilic cell-cell adhesion.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
 CC -!- TISSUE SPECIFICITY: Expressed in thymocytes and thymic stromal  
 CC cells; expression elevated in some T cell leukemias.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AF030455; AAC39762.1; -  
 CC EMBL; AF275945; AAF87240.1; -  
 CC EMBL; AF304447; AAG21183.1; -  
 CC EMBL; BC017774; AAHL1774.1; -  
 CC HSSP; P06907; 1NEU  
 CC Genew; HGNC:3496; EVAL  
 CC MIM; 604873; -  
 CC GO; GO:0005856; C:cytoskeleton; TAS.  
 CC GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.  
 CC GO; GO:0007156; P:homophilic cell adhesion; TAS.

DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003066; IG\_MHC.  
 DR InterPro: IPR003596; IG\_V.  
 DR InterPro: IPR000920; Myelin\_P0.  
 DR Pfam: PF00047; IG; 1.  
 DR PRINTS; PR00213; MYELINP0.  
 DR SMART; SM00406; IG; 1.  
 DR PROSITE; PS00835; IG LIKE; 1.  
 KW Cell adhesion; Immunoglobulin domain; Transmembrane; Glycoprotein;  
 KW Signal.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 215 EPITHELIAL V-LIKE ANTIGEN 1.  
 FT DOMAIN 27 154 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 155 175 POTENTIAL.  
 FT DOMAIN 176 215 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 27 141 IG-LIKE V-TYPE.  
 FT DISULFID 47 123 BY SIMILARITY.  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 215 AA; 24484 MW; 9D19EBA79B5815A CRC64;  
 Query Match 13.1%; Score 147; DB 1; Length 215;  
 Best Local Similarity 25.7%; Pred. No. 3.7e-06;  
 Matches 57; Conservative 35; Mismatches 80; Indels 50; Gaps 9;  
 Qy 7 LFPLASLVLYVSVCFVCVETAEVQGNPKLRCISCCKKEEVEATVVEVFRP 66  
 Db 21 LNPDA--VEIYTSRV-----LEAVNGTDARLACTSSSPAPVGDALTVT-WNFRP 67  
 Qy 67 -EGG-KDFLYIYVRNGHQEVESFFQRLQWNGSKDQVSTVNLVNLDSGLYTCNVSR 124  
 Db 68 LDGSPQFVFFVYHIDPFQMSGRFQKRVSDGNDNPERYDASILLWKLQDFDNGTTCQVKN 127  
 Qy 125 EPFEAHRFPVTKTRILRLPRTVEAGEDEFTSV-----SEIMWYLLV-----FLTLWL 173  
 Db 128 -----PDVGVGTEIRLSVTVTRFSEIHFLALAGSCALMIITVI 170  
 Qy 174 LIEMICYRVKSKAEPAQENASDYLAIPSENKENSAPVVEE 215  
 Db 171 VVVLQHYRKKQWAEKRA-----HKVVEIKSKEERLNQEK 205  
 RESULT 12  
 MYP0\_HETFR STANDARD; PRT; 246 AA.  
 ID MYP0\_HETFR  
 AC P20938;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Myelin P0 protein precursor (Myelin protein zero) (Myelin peripheral  
 DE protein) (MPP).  
 OS Heterodontus francisci (Horn shark).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;  
 OC Heterodontidae; Heterodontus.  
 OX NCBI\_TaxID=7792;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX TISSUE=Brain;  
 RX MEDLINE=90040744; PubMed=2478717;  
 RA Saavedra R.A., Fors L., Aebersold R.H., Arden B., Horvath S.,  
 RA Sanders J., Hood L.;  
 RT "The myelin proteins of the shark brain are similar to the myelin  
 RL J. Mol. Biol. 29:149-156(1989)  
 CC -!- FUNCTION: CREATION OF AN EXTRACELLULAR MEMBRANE PACE WHICH GUIDES  
 CC THE WRAPPING PROCESS AND ULTIMATELY COMPACTS ADJACENT LAMELLAE.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: FOUND ONLY IN PERIPHERAL NERVOUS SYSTEM  
 CC SCHWANN CELLS.  
 CC -!- PTM: N-GLYCANS IS SULFATED (BY SIMILARITY).  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -!- SIMILARITY: BELONGS TO THE MYELIN P0 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X16714; CAB37845.1; --  
 DR PIR; A32999; A32999.  
 DR HSSP; P06907; INEU.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003536; Ig V.  
 DR InterPro; IPR000920; Myelin\_P0.  
 DR PRINTS; PR00213; MYELINP0.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG-LIKE; 1.  
 DR PROSITE; PS00568; MYELIN\_P0; 1.  
 KW Myelin; structural protein; Glycoprotein; Transmembrane;  
 KW Phosphorylation; Immunoglobulin domain; Signal.  
 FT SIGNAL 1 27 BY SIMILARITY.  
 FT CHAIN 28 246 MYELIN P0 PROTEIN.  
 FT DOMAIN 28 150 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 151 178 POTENTIAL.  
 FT DOMAIN 179 246 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 28 143 IG-LIKE V-TYPE.  
 FT DISULFID 48 125 POTENTIAL.  
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. .) (COMPLEX) (BY  
 FT SIMILARITY).  
 SQ SEQUENCE 246 AA; 27335 NM; A776A9ED9D430FA0 CRC64;

Query Match 12.9%; Score 144.5; DB 1; Length 246;  
 Best Local Similarity 25.8%; Pred. No. 7.2e-06;  
 Matches 49; Conservative 37; Mismatches 77; Indels 27; Gaps 7;

QY 16 IYVSVCFVPCVVEPSETEAVQ-----GNPKLRICSCMKREVEATTVVVEFYRP 66  
 DB 10 LFCCSVLXAFSLVRPSQGISVSTHNLHKTGSDVLYC-GPWSNEYVSDLTLSWFRP 68  
 QY 67 EGKDFL-IYEVNQHVE--SPQGRLOWNSKLDQVSTVTLNLSGLYTCNV 123  
 DB 69 DNRDIISIFHYGVNVPYTERQGFGRVWVGDISKHDGSIIVNLDYIDNGFTCDVK 128  
 QY 124 REFEFAHRFPVKTTLPLRVTEE-----AGEDFTSVSEIMVILLVFLTLLEMI 178  
 DB 129 NPED-----VVGTSDDVHLTVYDKIPVVGAGVVSGLIGTFLGILLIVGLYL--FR 179  
 QY 179 YCVKVKAK 188  
 DB 180 YIVRRRARSE 189

RESULT 13  
 MYPO\_CHICK STANDARD; PRT; 249 AA.  
 ID MYPO\_CHICK  
 AC F37301;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Myelin P0 protein precursor (Myelin protein zero) (Myelin peripheral  
 DE protein) (MPP)  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90204597; PubMed=1690817;  
 RA Barbu M.;  
 RT "Molecular cloning of cDNAs that encode the chicken P0 protein:  
 RT evidence for early expression in avians."

RL J. Neurosci. Res. 25:143-151(1990).  
 CC -!- FUNCTION: CREATION OF AN EXTRACELLULAR MEMBRANE FACE WHICH GUIDES  
 CC THE WRAPPING PROCESS AND ULTIMATELY COMPACTS ADJACENT LAMELLAE.  
 CC -!- SUBCELLULAR LOCATION: type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: FOUND ONLY IN PERIPHERAL NERVOUS SYSTEM  
 CC SCHWANN CELLS.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -!- SIMILARITY: BELONGS TO THE MYELIN P0 PROTEIN FAMILY.  
 DR PIR; A61087; A61087.  
 DR HSSP; P06907; INEU.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003536; Ig V.  
 DR InterPro; IPR000920; Myelin\_P0.  
 DR PRINTS; PR00213; MYELINP0.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG-LIKE; 1.  
 DR PROSITE; PS00568; MYELIN\_P0; 1.  
 KW Myelin; structural protein; Glycoprotein; Transmembrane;  
 KW Phosphorylation; Immunoglobulin domain; Signal.  
 FT SIGNAL 1 29 BY SIMILARITY.  
 FT CHAIN 30 249 MYELIN P0 PROTEIN.  
 FT DOMAIN 30 153 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 154 179 POTENTIAL.  
 FT DOMAIN 180 249 CYTOPLASMIC.  
 FT DOMAIN 30 143 IG-LIKE V-TYPE.  
 FT DISULFID 50 127 POTENTIAL.  
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 249 AA; 27466 NM; FBD14801F8A08FB CRC64;

Query Match 12.5%; Score 140.5; DB 1; Length 249;  
 Best Local Similarity 23.6%; Pred. No. 3.7e-05;  
 Matches 48; Conservative 43; Mismatches 87; Indels 25; Gaps 8;

QY 13 LVLIVVSVCFP-----VCVEVPSSETEAVQGNPKLRICSCMKREVEATTVVVEFYRP 67  
 DB 13 LLLVGLLSASGSPSLAIHVTPREVYGVTVGSHVTLSC-SFWSSEWISDLSYTHWFAQE 71  
 QY 68 GKDFL-IYEVNQHVE--SPQGRLOWNSKLDQVSTVTLNLSGLYTCNVSR 124  
 DB 72 GSRDSISIFHYGVNVPYTERQGFGRVWVGDISKHDGSIIVNLDYIDNGFTCDVK 131  
 QY 125 REFEFAHRFPVKTTLPLRVTEEAGEDFTSVSEIMVILLVFLTLLEMI-YCY-- 181  
 DB 132 PPD-----IVGKSSQVTLVLEKVPTRYGVVGLSIGVLLVALLVAVVLYVFCWLR 185  
 QY 182 -----RKVSKABEA-AQENASD 197  
 DB 186 ROAVLQRLSAMKGLQRSAXD 208

RESULT 14  
 CIB2\_RAT STANDARD; PRT; 215 AA.  
 ID CIB2\_RAT  
 AC P54900;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Sodium channel beta-2 subunit precursor.  
 DE SCN2B.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-58; 95-134 AND 142-151.  
 RX MEDLINE=96067641; PubMed=8521473;  
 RA Isom L.L.; Ragsdale D.S.; de Jongh K.S.; Westenbroek R.E.;  
 RA Reber B.F.X.; Scheuer T.; Catterall W.A.;  
 RT "Structure and function of the beta 2 subunit of brain sodium  
 RT channel, a transmembrane glycoprotein with a CAM motif."  
 RL Cell 83:433-442(1995).

CC -!- FUNCTION: CRUCIAL IN THE ASSEMBLY, EXPRESSION, AND FUNCTIONAL  
 CC MODULATION OF THE HETEROTRIMERIC COMPLEX OF THE SODIUM CHANNEL.  
 CC THE BETA-2 SUBUNIT CAUSES AN INCREASE IN THE PLASMA MEMBRANE  
 CC SURFACE AREA AND IN ITS FOLDING INTO MICROVILLI.  
 CC -!- SUBUNIT: THE SODIUM CHANNEL CONSISTS OF A PORE-FORMING ALPHA  
 CC SUBUNIT, BETA-1 AND BETA-2 SUBUNITS. BETA-1 IS NONCOVALENTLY  
 CC ASSOCIATED WITH ALPHA, WHILE BETA-2 IS COVALENTLY LINKED BY  
 CC DISULFIDE BONDS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Brain-specific.  
 CC -!- DEVELOPMENTAL STAGE: DETECTED IN THE EARLIEST PHASE OF  
 CC NEUROGENESIS IN BRAIN, AND EXPRESSION IS GREATLY INCREASED  
 CC CONCOMITANT WITH AXON EXTENSION AND SYNAPTONEGENESIS.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 CC EMBL; U37026; AAC52967.1; -.  
 CC EIR; AS7843; AS7843. -.  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003599; Ig.  
 CC InterPro; IPR000920; Myelin\_P0.  
 CC PRINTS; PR00213; MYELIN\_P0.  
 CC SMART; SM00409; IG; 1.  
 CC PROSITE; PS50835; IG\_LIKE; 1.  
 CC Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;  
 KW Glycoprotein; Signal; Immunoglobulin domain.  
 FT SIGNAL 1 29  
 FT CHAIN 30 215 SODIUM CHANNEL BETA-2 SUBUNIT.  
 FT DOMAIN 30 159 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 160 180 POTENTIAL.  
 FT DOMAIN 181 215 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 32 154 IG-LIKE C2-TYPE.  
 FT DISULFID 55 127 POTENTIAL.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 80 80 F -> V.  
 SQ SEQUENCE 215 AA; 24144 MW; 23D432EA83C1459D CRC64;  
 Query Match 10.5%; Score 118; DB 1; Length 215;  
 Best Local Similarity 25.1%; Pred. No. 0.0015;  
 Matches 49; Conservative 41; Mismatches 85; Indels 20; Gaps 8;  
 QY 27 VEVPSFEAVQGNPKMLRCI--SCMKREVEATTVVEVFPPEG--KDFLIYEVNKH 81  
 Db 32 VTPTTSLVLSGSDRLPCTFNFCYTVNKHQFS--LAWTYQECNCEEMFLOFRMKIIN 89  
 QY 82 QEVSPFQGLQWNGSKDLQDVSTIVLVNLDGLYTCNVSRFEFEAEHPFVKTRLI 141  
 Db 90 LKLER-FGRVFEFGNSPKYDVSIVLKNVQLEDEGIYCYITN--PDRHGHCK--I 142  
 QY 142 PLRVTEAGEDEFTSVSEIMMYLLVFLVTLMLLIEMIV-CYKVKSKAEBAQAENASYLA 200  
 Db 143 YLQVLLEVPFERDSTVAVIVGASVGGLAVILVILVWVKVRR-----KKEQKLSTDDLK 197  
 QY 201 IPSENKENSAPVEE 215  
 Db 198 TEEGKTDGEHNAED 212  
 RESULT 15  
 FGR2 MOUSE  
 ID\_FGR2\_MOUSE STANDARD; PRT; 821 AA.  
 AC P21803; O55141; Q00389; Q61342;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Fibroblast growth factor receptor 2 precursor (EC 2.7.1.112) (FGFR-2)  
 DE (Keratinocyte growth factor receptor).  
 DE FGR2 OR ECT1 OR BEK.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RC TISSUE=Brain, and Liver;  
 RX MEDLINE=92228773; PubMed=1373495;  
 RA Manukhani A., Dell'Era P., Moscatelli D., Kornbluth S.,  
 RA Hanafusa H., Basilico C.;  
 RT "Characterization of the murine BEK fibroblast growth factor (FGF)  
 RT receptor: activation by three members of the FGF family and  
 RT requirement for heparin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:3305-3309(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RC STRAIN=BAUB/C; TISSUE=Brain;  
 RX MEDLINE=91270892; PubMed=1711190;  
 RA Raz V., Kelman Z., Avioli A., Neufeld G., Givol D., Yarden Y.;  
 RT "PCR-based identification of new receptors: molecular cloning of a  
 RT receptor for fibroblast growth factors.";  
 RL Oncogene 6:753-760(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RX MEDLINE=91095977; PubMed=1846048;  
 RA Miki T., Fleming T.P., Bottaro D.P., Rubin J.S., Ron D.,  
 RA Aaronson S.A.;  
 RT "Expression cDNA cloning of the KGF receptor by creation of a  
 RT transforming autoclone loop.";  
 RL Science 251:72-75(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=98167854; PubMed=9499422;  
 RA Wigg S.R.F., Burns H.D., Oldridge M., Heath J.K., Wilkie A.O.M.;  
 RT "Conserved use of non-canonical 5' splice site (5' splice site) in alternative  
 RT splicing by fibroblast growth factor receptors 1, 2 and 3.";  
 RL Hum. Mol. Genet. 7:685-691(1998).  
 RN [5]  
 RP SEQUENCE OF 477-821 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=89219016; PubMed=2468999;  
 RA Kornbluth S., Paulson X.E., Hanafusa H.;  
 RT "Novel tyrosine kinase identified by phosphotyrosine antibody  
 RT screening of cDNA libraries.";  
 RL Mol. Cell Biol. 8:5541-5544(1988).  
 CC -!- FUNCTION: RECEPTOR FOR ACIDIC AND BASIC FIBROBLAST GROWTH FACTORS.  
 CC POSSESSES A HIGHER AFFINITY FOR ACIDIC THAN FOR BASIC FGF'S.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Long;  
 CC IsoId=P21803-1; Sequence=Displayed;  
 CC Name=Short;  
 CC IsoId=P21803-2; Sequence=VSP 002985, VSP 002986, VSP 002987;  
 CC -!- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR  
 CC FAMILY.  
 CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 CC EMBL; M86441; AAA37286.1; -.

DR EMBL; X55441; CAA39083.1; --  
 DR ENBL; M62503; AAA39377.1; --  
 DR ENBL; Y16152; CAA66098.1; --  
 DR ENBL; Y16167; CAA66099.1; --  
 DR ENBL; M23362; AAA37285.1; --  
 DR PIR; A44142; TMSBK.  
 DR PIR; S17295; S17295.  
 DR HSP; P11362; LFGL.  
 DR MGD; M19523; Fgfr2.  
 DR GO; GO:0007435; Pissalivary gland morphogenesis; IMP.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_C2.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00047; ig; 3.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00408; Igc2; 3.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS0835; IG\_LIKS; 3.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;  
 KW Transferase; Phosphorylation; Transmembrane; Signal;  
 KW Immunoglobulin domain; Repeat; Alternative splicing.  
 FT SIGNAL 1 21  
 FT CHAIN 22 821 FIBROBLAST GROWTH FACTOR RECEPTOR 2.  
 FT DOMAIN 22 377 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 378 398 POTENTIAL.  
 FT DOMAIN 399 821 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 39 125 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 154 247 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 256 358 IG-LIKE C2-TYPE 3.  
 FT DOMAIN 481 770 PROTEIN KINASE.  
 FT NP\_BIND 487 495 ATP (BY SIMILARITY).  
 FT BINDING 517 517 ATP (BY SIMILARITY).  
 FT ACT\_SITE 626 626 BY SIMILARITY.  
 FT MOD\_RES 657 657 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT DISULFID 62 107 POTENTIAL.  
 FT DISULFID 179 231 POTENTIAL.  
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 123 123 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARSPLIC 37 37 E -> G (in isoform Short).  
 FT VARSPLIC 38 152 Missing (in isoform Short).  
 FT VARSPLIC 314 361 /FTid=VSP\_002986  
 AACWNTDKIEVLIRNVTEDAGEYTCLAGNSIGISFHS  
 AMTLVLP -> HSGINSSNAEVLALFNVTEDAGEYICKVS  
 NYIGQANQSAMTLVLPKQK (in isoform Short).  
 /FTid=VSP\_002987  
 A -> V (IN REF. 2).  
 GE -> RG (IN REF. 2).  
 E -> R (IN REF. 2).  
 I -> Y (IN REF. 2).  
 DV -> R (IN REF. 2).  
 C -> V (IN REF. 2 AND 3).  
 S -> P (IN REF. 2 AND 3).  
 W -> R (IN REF. 2).  
 Y -> I (IN REF. 2).  
 E -> R (IN REF. 2).  
 N -> Y (IN REF. 2).  
 821 AA; 91983 MW; FCDE28ADDG61F4414 CRC64;  
 FT CONFLICT 53 53  
 FT CONFLICT 55 56  
 FT CONFLICT 90 90  
 FT CONFLICT 119 119  
 FT CONFLICT 142 143  
 FT CONFLICT 169 169  
 FT CONFLICT 187 187  
 FT CONFLICT 214 214  
 FT CONFLICT 229 229  
 FT CONFLICT 275 275  
 FT CONFLICT 301 301  
 FT SEQUENCE 821 AA; 91983 MW; FCDE28ADDG61F4414 CRC64;

Query Match 9.9%; Score 111.5; DB 1; Length 821;  
 Best Local Similarity 25.9%; Pred. No. 0.025; 55; Indels 39; Gaps 7;  
 Matches 41; Conservative 23; Mismatches 23; Indels 39; Gaps 7;  
 QY 83 EVESPFQGRIGW-----NGSK-----DLQDVSIIVL---NVTINDSG 116  
 DB 279 KVISDAQPHIQIKHVKESKNGDPGLPYLKVLAAGVNTTDEIEVLIRNVTEDAG 338  
 QY 117 LYTGNVSREFEPEAHPFVKTTTLPLRVTE-EAGEDFTSVVSEIMMVLILVLTLLI 175  
 DB 339 EYTCLAGNSIGISFHSAML-TVLPAPVREKEITASPDYL---EIAIYIGVFLIACMVV 393  
 QY 176 EMICYRVKVKABEAAQENASDYLAIPESENKENSAPV 213  
 DB 394 TVIFCRMKTTTK-----PDFSSQPAVHKLTKRIPL 424

Search completed: December 5, 2003, 23:19:06  
 Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 22:55:42 ; Search time 109 Seconds  
(without alignments)  
509.003 Million cell updates/sec

Title: US-09-936-680-2

Perfect score: 1124  
Sequence: 1 MPANRFLPLASLVLIWVS.....SDYLAIPSENKENSVPVEE 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1124	100.0	230	4 Q9ULR2	Q9ulr2 homo sapien
2	1116	99.3	215	6 Q8HXJ7	Q8hxj7 macaca fasc
3	1105	98.3	215	11 Q9JK00	Q9jk00 rattus norv
4	1097	97.6	215	11 Q8BHK2	Q8bhk2 mus musculu
5	1091	97.1	215	11 Q91299	Q91299 mus musculu
6	420	37.4	185	4 Q8WU42	Q8wu42 homo sapien
7	262.5	23.4	273	11 Q9QUX3	Q9qux3 rattus norv
8	178.5	15.9	203	13 Q8JFG6	Q8jfg6 brachydanio
9	163.5	14.5	251	4 Q14902	Q14902 homo sapien
10	163.5	14.5	258	4 Q9BR67	Q9br67 homo sapien
11	160.5	14.3	225	4 Q92677	Q92677 homo sapien
12	150	13.3	202	13 Q91406	Q91406 salmo sp. i
13	147	13.1	215	11 Q91WT4	Q91wt4 mus musculu
14	147	13.1	229	13 Q8AVM3	Q8avm3 xenopus lae
15	131	11.7	243	4 Q9UELA	Q9uel4 homo sapien
16	131	11.7	269	4 Q95297	Q95297 homo sapien

Q9uel6 homo sapien  
Q9nyk4 homo sapien  
Q8jg36 brachydanio  
Q8ix11 homo sapien  
Q8ix39 homo sapien  
Q9d7b8 mus musculu  
Q62861 rattus norv  
Q8ix38 homo sapien  
Q8k190 rattus norv  
Q8r373 mus musculu  
Q9tct3 oryctolagus  
Q920e5 mus musculu  
Q8n917 homo sapien  
Q99052 mus musculu  
Q9h6b4 homo sapien  
Q61563 mus musculu  
Q96m70 homo sapien  
Q96kv6 homo sapien  
Q9x293 rattus norv  
Q63241 rattus norv  
Q63242 rattus norv  
Q31406 gallus galli  
Q63237 rattus norv  
Q8nf70 homo sapien  
Q8uwk1 ictalurus p  
Q8uwk1 ictalurus p  
Q91wp1 mus musculu  
Q8k094 mus musculu  
Q8bvf6 mus musculu

# ALIGNMENTS

## RESULT 1

Q9ULR2 PRELIMINARY; PRT; 230 AA.

AC Q9ULR2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein KIAA1158 (Fragment).  
GN KIAA1158.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=20039618; PubMed=10574461;

RA Hirotsawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;

RT "Characterization of cDNA clones selected by the GeneMark analysis

RL DNA Res. 6:329-336(1999).

DR EMBL; AB032984; BAB86472.1; -

DR HSSP; P06907; INEU.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003006; IG\_MHC.

DR Pfam; PF00047; IG; 1.

DR SMART; SM00409; IG; 1.

DR PROSITE; PS50835; IG LIKE; 1.

KW Hypothetical protein.

FT NON\_TER 1

SQ SEQUENCE 230 AA; 26357 MW; B06D5155E5FE5F98 CRC64;

Query Match

Best Local Similarity 100.0%; Score 1124; DB 4; Length 230;

Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MPANRFLPLASLVLIWVSVCPPVCVEVSETEAVOGNPMKLRICISCMREVEATTVV 60

Db 16 MPAFNRPLASLVLIYVSVCFVCPVSETEAVQGNPKMLRCISCMKREEVEATTVV 75  
 QY 61 EMFYRPEGKDFLIYERNGHQEVESPPQGLQWNGSKLDQVSIITVNTLNDGLTYC 120  
 Db 76 EMFYRPEGKDFLIYERNGHQEVESPPQGLQWNGSKLDQVSIITVNTLNDGLTYC 135  
 QY 121 NVSRFEFEARPPFKVTKTRLLPLRVTEAGDFTSVVSEIMMYILLVFLTLMLLIEMTYC 180  
 Db 136 NVSRFEFEARPPFKVTKTRLLPLRVTEAGDFTSVVSEIMMYILLVFLTLMLLIEMTYC 195  
 QY 181 YRKVSKABEAAQENASDYLAIPSENKENSAPVVEE 215  
 Db 196 YRKVSKABEAAQENASDYLAIPSENKENSAPVVEE 230

RESULT 2  
 Q8HXJ7 PRELIMINARY; PRT; 215 AA.  
 ID Q8HXJ7  
 AC Q8HXJ7  
 DT 01-MAR-2003 (TremBLrel. 23, Created)  
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
 DE Hypothetical protein.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Medulla oblongata;  
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Sugano S.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Medulla oblongata;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirata M., Suto Y.,  
 RA Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Assignment of 118 novel cDNAs of cynomolgus monkey brain to human  
 RT chromosomes.";  
 RL Gene 275:31-37(2001).  
 DR EMBL; AB097521; BAC41746.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 215 AA; 24702 MW; 25319D5ED218AACF CRC64;

Query Match 99.3%; Score 1116; DB 6; Length 215;  
 Best Local Similarity 99.1%; Pred. No. 6.8e-109;  
 Matches 213; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MPAFNRPLASLVLIYVSVCFVCPVSETEAVQGNPKMLRCISCMKREEVEATTVV 60  
 Db 1 MPAFNRPLASLVLIYVSVCFVCPVSETEAVQGNPKMLRCISCMKREEVEATTVV 60  
 QY 61 EMFYRPEGKDFLIYERNGHQEVESPPQGLQWNGSKLDQVSIITVNTLNDGLTYC 120  
 Db 61 EMFYRPEGKDFLIYERNGHQEVESPPQGLQWNGSKLDQVSIITVNTLNDGLTYC 120  
 QY 121 NVSRFEFEARPPFKVTKTRLLPLRVTEAGDFTSVVSEIMMYILLVFLTLMLLIEMTYC 180  
 Db 121 NVSRFEFEARPPFKVTKTRLLPLRVTEAGDFTSVVSEIMMYILLVFLTLMLLIEMTYC 180  
 QY 181 YRKVSKABEAAQENASDYLAIPSENKENSAPVVEE 215  
 Db 181 YRKVSKABEAAQENASDYLAIPSENKENSAPVVEE 215

RESULT 3  
 Q9JK00 PRELIMINARY; PRT; 215 AA.  
 ID Q9JK00  
 AC Q9JK00  
 DT 01-OCT-2000 (TremBLrel. 15, Created)  
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
 DE Voltage-gated sodium channel beta-3 subunit (Sodium channel beta 3 subunit).  
 GN SCN3B OR SCN3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20160948; PubMed=10688874;  
 RA Morgan K., Stevens E.B., Shaw B., Cox P., Dixon A.K., Lee K.,  
 RA Pincock R.D., Hughes J., Richardson P.J., Mizuguchi K., Jackson A.P.;  
 RT "beta3: An additional auxiliary subunit of the voltage-sensitive  
 RT sodium channel that modulates channel gating with distinct kinetics.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2308-2313(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Dorsal root ganglion;  
 RA Qu Y., Curtis R., Lawson D., Gilbride K., Ge P., DiStefano P.S.,  
 RA Silos-Santiago I., Catterall W.A., Scheuer T.;  
 RT "Differential Modulation of Sodium Channel Gating and Persistent  
 RT Sodium Currents by the beta1, beta2, and beta3 Subunits.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ243395; CAB76838.1; -.  
 DR EMBL; AF378093; AAKS5415.1; -.  
 DR HSSP; P06907; INBU.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS0835; IG\_LIKE; 1.  
 KW Ionic channel.  
 SQ SEQUENCE 215 AA; 24799 MW; 056B488EF5AEAF4F CRC64;

Query Match 98.3%; Score 1105; DB 11; Length 215;  
 Best Local Similarity 98.1%; Pred. No. 9.7e-108;  
 Matches 211; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MPAFNRPLASLVLIYVSVCFVCPVSETEAVQGNPKMLRCISCMKREEVEATTVV 60  
 Db 1 MPAFNRPLASLVLIYVSVCFVCPVSETEAVQGNPKMLRCISCMKREEVEATTVV 60  
 QY 61 EMFYRPEGKDFLIYERNGHQEVESPPQGLQWNGSKLDQVSIITVNTLNDGLTYC 120  
 Db 61 EMFYRPEGKDFLIYERNGHQEVESPPQGLQWNGSKLDQVSIITVNTLNDGLTYC 120  
 QY 121 NVSRFEFEARPPFKVTKTRLLPLRVTEAGDFTSVVSEIMMYILLVFLTLMLLIEMTYC 180  
 Db 121 NVSRFEFEARPPFKVTKTRLLPLRVTEAGDFTSVVSEIMMYILLVFLTLMLLIEMTYC 180  
 QY 181 YRKVSKABEAAQENASDYLAIPSENKENSAPVVEE 215  
 Db 181 YRKVSKABEAAQENASDYLAIPSENKENSAPVVEE 215

RESULT 4  
 Q8BHK2 PRELIMINARY; PRT; 215 AA.  
 ID Q8BHK2  
 AC Q8BHK2  
 DT 01-MAR-2003 (TremBLrel. 23, Created)  
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
 DE Voltage-gated sodium channel beta-3 subunit.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head, and Spinal cord;



```

DR Pfam: PF00047; ig; 1.
SW Ionic channel.
KQ SEQUENCE 273 AA; 31063 MW; 156A31899A906849 CRC64;

Query Match      23.4%; Score 262.5; DB 11; Length 273;
Best local similarity 39.3%; Pred. No. 3.5e-19;
Matches 68; Conservative 23; Mismatches 53; Indels 29; Gaps 5;

QY      10 LASLVLYWVSVCFFVPCVEVSETEAVQGPMKLCISCKMKEEVEATTVWEFVRPGG 69
DB      5 LALVGVANLWSAGGCVDESETEAYVTGMFTFLICLSCKRSETTAEFTFWPRQGT 64
QY      70 KDFL-IYEYRNHGQEVESP--FOGRLOWMS---KOLQPVSITVLNLTNSGLTYCNVS 123
DB      65 EEFVKILRYENEVLQEEDERFEGRVVWGSRGKTQLQLSIFITNTVYNHSGDYECVY 124
QY      124 REFEAPRPFKVTKTRLPLRVTEAGEDFTSVVSEIMMYIL----LVFLTL 172
DB      125 RLUFFDNY-----EHNTSVVKIHLDEVOKGWSLVLLM 158

RESULT 8
Q8JFG6 PRELIMINARY; PRT; 203 AA.
ID OSJFG6 AC OSJFG6;
AC AC OSJFG6;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Myelin protein zero.
GN PO.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
[1]
FN PROSITE; PS50835; IG LIKE; 1.
RP SEQUENCE FROM N.A.
RC TSSUE=Myelin;
RA Schweitzer J., Becker T., Becker C.G., Schachner M.;
RT "Increased expression of protein zero (PO) in the CNS during
RT regeneration of axons in adult zebrafish.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR ENML: AJ489219; CAD32961.1; --
DR InterPro; IPR003599; IG-like.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR InterPro; IPR000920; Myelin_PO.
DR Pfam; PF00047; ig; 1.
DR PRINTS; PR00213; MYELINPO.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 203 AA; 22097 MW; 12614E9076D373D1 CRC64;

Query Match      15.9%; Score 178.5; DB 13; Length 203;
Best Local Similarity 26.9%; Pred. No. 1.6e-10;
Matches 53; Conservative 41; Mismatches 86; Indels 17; Gaps 7;

QY      10 LASLVLYWVSVCFFVPCVEVSETEAVQGPMKLCISCKMKEEVEATTVWEFVRPGG 69
DB      7 LSLVLLGTASSTALVNVTUSEKIALVGSDVRLSC-SFFSNQWTSPFSFTWHRRPDGA 65
QY      70 KDFL-IYEYRNHGQEVESPFOGRLOWMSKOLQDVSIIVNLVNTNSGLTYCNVSREF 126
DB      66 KDAISIFHYGGGEAYPANKGFQNRLEFPVGNPSRDGSILLKNLFDNGTGTFCTCAKNPP 125
QY      127 EPEAPRPFKVTKTRLPL-RVTEAGEDFTSVVSEIMMYILV---FLTLLLIEMVYC- 180
DB      126 DIGCH--PSTRLLVFVKPVQAGVTGSIIGVIGLLILVVAIYYLMRFIVARRVFLS 182
QY      181 ----YRKVSKAEAAQE 193

```



01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 Epithelial V-like antigen.  
 EVA.  
 Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RA Strausberg R.;  
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mesonephros;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 DR EMBL; BC015076; AAH15076.1; -;  
 DR EMBL; AK032768; BAC28022.1; -;  
 DR MGD; MGI:1289160; Eva.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR InterPro; IPR000920; Myelin\_p0.  
 DR Pfam; PF00047; Ig; 1.  
 DR PRINTS; PR00213; MYELINP0.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 215 AA; 24092 MW; F6E5E36787CF79C4 CRC64;  
 Query Match 13.1%; Score 147; DB 11; Length 215;  
 Best Local Similarity 23.4%; Pred. No. 3.5e-07;  
 Matches 47; Conservative 45; Mismatches 101; Indels 8; Gaps 5;  
 QY 10 LASLVLYVWVCFVVCVEV--PSETEAVQGNPKLRCISCMKREVEATTVEVFRP- 66  
 DB 10 LPLLSLQTLALCTPAVEIYTSGALEAVNGTDVRUKCTFSSGAPVGDALTVT--WNRFR 68  
 QY 67 EGKGD-FLIYEVNANGHEVESPFQGRLOWNGSKLDQVSVITVLNVTINDSGLYTCNVSR 125  
 DB 69 DGGREQEVFYHMDPFPMSGRKDRVWDGPNERYDVSVILLWKLQFDNDGTYTCQKNP 128  
 QY 126 FFEARHPVKTRLIPLRVTEAGEDTSVV-SIMMYILLVFLTLMLLIEMIYC----- 185  
 DB 129 PDVGG---IVGTIRLSVHTVPFSEIYFLAVAIGSACALMIIVVWVLFQHFRRKRWAD 185  
 QY 186 KAEAAQENASDYLAIIPSENK 206  
 DB 186 SADKAEGTKSKEELNQGNK 206  
 RESULT 14  
 Q8AVM3  
 ID Q8AVM3 PRELIMINARY; PRT; 229 AA.  
 AC Q8AVM3;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodidae; Xenopus.  
 NCBI\_TaxID=8355;  
 [1]  
 SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC041735; AAH41735.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 229 AA; 25865 MW; 986A3F14B8261665 CRC64;  
 Query Match 13.1%; Score 147; DB 13; Length 229;  
 Best Local Similarity 26.8%; Pred. No. 3.8e-07;  
 Matches 57; Conservative 35; Mismatches 93; Indels 28; Gaps 9;  
 QY 14 VLIYVWVCFVVCVEVSETEA--VOGNPKLRCISCMKREVEATTVEVFRP-EGG 69  
 DB 3 VLIILLAGPPVLCIDISHHPDAFGVWGSKILKC-SFTSSYPISDIVAVDVTYRQLNGG 61  
 QY 70 KDFLIYEVN-NGHEVESPFQGRLOWNGSKLDQVSVITVLNVTINDSGLYTCNVSR 128  
 DB 62 STVTIILHFQKPYPILEGPFKDRIIWEGDVRGDASISLTLRLTNGTLSCIVWNP 121  
 QY 129 EAHRRPVKTRLIPLRVTEAGEDTSVV-SIMMYILLVFLTLMLLIEMIYC----- 180  
 DB 122 HGNVQTKLTVI-----ENLFFQNTVILLSALVFIPSAVLSLLILMRRSINRGRTK 176  
 QY 181 ---YRKVSKAEAA-----QENASDYLAIIPSE 204  
 DB 177 NLKWRKKSPIESQCVVDDNENTPLHPTLPQE 209  
 RESULT 15  
 Q9UELA  
 ID Q9UELA PRELIMINARY; PRT; 243 AA.  
 AC Q9UELA;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 RA Rhodes S.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL035308; CAA22913.1; -;  
 DR HSSP; P06907; INEU.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig MHC.  
 DR InterPro; IPR003596; Ig V.  
 DR InterPro; IPR000920; Myelin\_p0.  
 DR Pfam; PF00047; Ig; 1.  
 DR PRINTS; PR00213; MYELINP0.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 243 AA; 26551 MW; DF4E85BE8D3C005D CRC64;  
 Query Match 11.7%; Score 131; DB 4; Length 243;  
 Best Local Similarity 23.0%; Pred. No. 1.9e-05;  
 Matches 51; Conservative 32; Mismatches 81; Indels 58; Gaps 9;  
 QY 27 VEVPSETEAVQGNPKLRCISCMKREVEAT---TVVWFVYRPGGKDFL-IYEVNCHQ 82  
 DB 14 VYTPKEIFVANGTQKLTG---KFKSTSTTGITSVNSFQEGADTTVSFHYSGQV 69  
 QY 83 EVES--PFQGRLOWNGSKLDQVSVITVLNVTINDSGLYTCNVSRPFVFAHRPVKTTRL 140  
 DB 70 YLGNYPFKDRISWAGLDKKDASINENMQFHNGTYICDVKN----- 113  
 QY 141 ILELVTEAGEDTSVVSE-----IMMYILLVFLTLMLLIEMIYC---YRK----- 183  
 DB 114 -PDIVVQGHIRLYVVEKENLNPVFWVWVIVTAVVLGLTLLISILAVLYRRKNSKR 172  
 QY 184 -----VSKAEAAQENASD-----YLAIPSENKENSANV 211

```

Db      173 DYTGCSTSESLSPVKAPRKSPDTEGLVKSLPSGSHQGPVI 214
          :! ::! ::! ::! ::! ::! ::! ::! ::! ::! ::! ::! ::! ::! ::! ::!

```

Search completed: December 5, 2003, 23:22:33  
Job time : 112 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 23:17:13 ; Search time 30 Seconds  
(without alignments)  
303.228 Million cell updates/sec

Title: US-09-936-680-2  
Perfect score: 1124  
Sequence: 1 MPAPNRLPLASLVLYRWSS.....SDYLAIPSENKENSAPVVEE 215

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pdp.\*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pdp.\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pdp.\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pdp.\*  
5: /cgn2\_6/ptodata/1/1aa/ECTUS COMB.pdp.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	13.1	215	4	US-09-996-243-389
2	131	11.7	269	4	US-09-430-503-2
3	131	11.7	269	4	US-09-430-503-4
4	131	11.7	269	4	US-09-430-503-6
5	131	11.7	269	4	US-09-430-503-8
6	129	11.5	199	4	US-09-430-503-44
7	129	11.5	199	4	US-09-430-503-48
8	127	11.3	199	4	US-09-430-503-42
9	127	11.3	199	4	US-09-430-503-46
10	127	11.3	209	4	US-09-430-503-18
11	127	11.3	209	4	US-09-430-503-20
12	127	11.3	209	4	US-09-430-503-22
13	127	11.3	209	4	US-09-430-503-24
14	125	11.1	270	4	US-09-430-503-26
15	125	11.1	270	4	US-09-430-503-28
16	125	11.1	270	4	US-09-430-503-30
17	125	11.1	270	4	US-09-430-503-32
18	118	10.5	159	4	US-09-430-503-34
19	118	10.5	159	4	US-09-430-503-36
20	118	10.5	159	4	US-09-430-503-38
21	118	10.5	159	4	US-09-430-503-40
22	112.5	10.0	318	4	US-09-656-952-2
23	109	9.7	344	4	US-09-656-952-19
24	109	9.7	394	4	US-09-656-952-20
25	109	9.7	394	4	US-09-996-243-422
26	108.5	9.7	373	4	US-09-996-243-503
27	108	9.6	801	3	US-09-383-630-6

Sequence 21, Appl  
Sequence 21, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 29, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 28, Appl  
Sequence 2, Appl  
Sequence 62, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 6, Appl

28 106 9.4 519 3 US-08-996-338-21  
29 106 9.4 519 4 US-09-556-972-21  
30 106 9.4 537 1 US-08-604-333-4  
31 106 9.4 537 3 US-09-110-618-4  
32 106 9.4 537 4 US-09-173-151A-29  
33 106 9.4 537 4 US-09-578-178-4  
34 106 9.4 537 4 US-09-577-806-4  
35 99.5 8.9 821 2 US-08-451-822A-13  
36 99.5 8.9 821 4 US-08-323-430-13  
37 99 8.8 541 1 US-08-604-333-2  
38 99 8.8 541 3 US-09-110-618-2  
39 99 8.8 541 4 US-09-173-151A-28  
40 99 8.8 541 4 US-09-578-178-2  
41 99 8.8 541 4 US-09-577-806-2  
42 98.5 8.8 408 4 US-09-724-864-62  
43 98.5 8.8 521 3 US-08-996-338-20  
44 98.5 8.8 521 4 US-09-556-972-20  
45 98.5 8.8 729 1 US-08-070-165F-6

## ALIGNMENTS

## RESULT 1

US-09-996-243-389  
; Sequence 389, Application US/09996243  
; Patent No. 6478825

## GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC13  
; CURRENT APPLICATION NUMBER: US/09/996,243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28





RESULT 5  
US-09-430-503-8  
; Sequence 8, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-8

Query Match 11.7%; Score 131; DB 4; Length 269;  
Best Local Similarity 23.0%; Pred. No. 4.4e-06;  
Matches 51; Conservative 32; Mismatches 81; Indels 58; Gaps 9;

QY 27 VEVPSETEAVQGNPMKRLCISCCKREVEAT--TVVEWFYRPEGKDFL-IYEYRNHGQ 82  
Db 40 VYTPKEIFVANGTQGLKTC-----KFKSTSTTGLTSVMSFQPEGADTTVSFFHYSQGV 95  
QY 83 EVES--PFQGRLOWNGSKDLQDVSIITLVNLTNDGLYTCNVSRFEFEFAHRPFVKTTTL 140  
Db 96 YLGNYPFKDRISWAGLDKDKASININMQFIHNGTYICDVKN-----139

QY 141 IPLRVTEAGEDFTSVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQPGHRIYVVEKENLPVFPVWVVGIVTAVLGLTLLISMLAVLYRRKNSKR 198

184 -----VSKAEAAQNASD-----YLAIPSENKENSAN 211  
199 DYTGCSTSESLSVPKQAPRSPDTEGLVLSLPSGSHQGPVI 240

RESULT 6  
US-09-430-503-44  
; Sequence 44, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 44  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-44

Query Match 11.5%; Score 129; DB 4; Length 199;  
Best Local Similarity 26.0%; Pred. No. 4.4e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;

QY 27 VEVPSETEAVQGNPMKRLCISCCKREVEAT--TVVEWFYRPEGKDFL-IYEYRNHGQ 82  
Db 40 VYTPKEIFVANGTQGLKTC-----KFKSTSTTGLTSVMSFQPEGADTTVSFFHYSQGV 95  
QY 83 EVES--PFQGRLOWNGSKDLQDVSIITLVNLTNDGLYTCNVSRFEFEFAHRPFVKTTTL 140  
Db 96 YLGNYPFKDRISWAGLDKDKASININMQFIHNGTYICDVKN-----139

QY 141 IPLRVTEAGEDFTSVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQPGHRIYVVEKENLPVFPVWVVGIVTAVLGLTLLISMLAVLYRRKNSKR 198

188 E 188  
199 D 199

RESULT 7  
US-09-430-503-48  
; Sequence 48, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-48

Query Match 11.5%; Score 129; DB 4; Length 199;  
Best Local Similarity 26.0%; Pred. No. 4.4e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;

QY 27 VEVPSETEAVQGNPMKRLCISCCKREVEAT--TVVEWFYRPEGKDFL-IYEYRNHGQ 82  
Db 40 VYTPKEIFVANGTQGLKTC-----KFKSTSTTGLTSVMSFQPEGADTTVSFFHYSQGV 95  
QY 83 EVES--PFQGRLOWNGSKDLQDVSIITLVNLTNDGLYTCNVSRFEFEFAHRPFVKTTTL 140  
Db 96 YLGNYPFKDRISWAGLDKDKASININMQFIHNGTYICDVKN-----139

QY 141 IPLRVTEAGEDFTSVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQPGHRIYVVEKENLPVFPVWVVGIVTAVLGLTLLISMLAVLYRRKNSKR 198

188 E 188  
199 D 199

RESULT 8  
US-09-430-503-42  
; Sequence 42, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 42  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-42

Query Match 11.3%; Score 127; DB 4; Length 199;  
Best Local Similarity 26.0%; Pred. No. 7.3e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;

QY 27 VEVPSETEAVQGNPMKRLCISCCKREVEAT--TVVEWFYRPEGKDFL-IYEYRNHGQ 82

QY 141 IPLRVTEAGEDFTSVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQPGHRIYVVEKENLPVFPVWVVGIVTAVLGLTLLISMLAVLYRRKNSKR 198

188 E 188  
199 D 199

RESULT 7  
US-09-430-503-48  
; Sequence 48, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-48

Query Match 11.5%; Score 129; DB 4; Length 199;  
Best Local Similarity 26.0%; Pred. No. 4.4e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;

QY 27 VEVPSETEAVQGNPMKRLCISCCKREVEAT--TVVEWFYRPEGKDFL-IYEYRNHGQ 82  
Db 40 VYTPKEIFVANGTQGLKTC-----KFKSTSTTGLTSVMSFQPEGADTTVSFFHYSQGV 95  
QY 83 EVES--PFQGRLOWNGSKDLQDVSIITLVNLTNDGLYTCNVSRFEFEFAHRPFVKTTTL 140  
Db 96 YLGNYPFKDRISWAGLDKDKASININMQFIHNGTYICDVKN-----139

QY 141 IPLRVTEAGEDFTSVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQPGHRIYVVEKENLPVFPVWVVGIVTAVLGLTLLISMLAVLYRRKNSKR 198

188 E 188  
199 D 199

RESULT 8  
US-09-430-503-42  
; Sequence 42, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 42  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-42

Query Match 11.3%; Score 127; DB 4; Length 199;  
Best Local Similarity 26.0%; Pred. No. 7.3e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;

QY 27 VEVPSETEAVQGNPMKRLCISCCKREVEAT--TVVEWFYRPEGKDFL-IYEYRNHGQ 82

Db 40 VYTPKEIFVANGTQGLKTC---KFKSTSTTGLTSVSWSPQEGADTTVSFFHYSQGV 95  
QY 83 EVES---PFQGRLOWNGSKDLQDVSTVLNVLTNDLSGLYTCNVSRFEFEHRRPFVKTTTL 140  
Db 96 YLGNYPFKRISWAGDLKKDASINENMQFIHNGTYICDVKN----- 139  
QY 141 IPLRVTEAGEDFTSVVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQGHIRLYVVEKENLPVFPVWVVGIVTAVVLGTLTLLSMILAVLYRKNKR 198  
QY 188 E 188  
Db 199 D 199

## RESULT 9

US-09-430-503-46  
; Sequence 46, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 46  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-46

Query Match 11.3%; Score 127; DB 4; Length 199;  
Best Local Similarity 26.0%; Pred. No. 7.3e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;  
QY 27 VEVPSETEAVQGNPKRLCISCMKREVEAT---TVVEWVYRPEGGKDFL-IYEYRNGHQ 82  
Db 40 VYTPKEIFVANGTQGLKTC---KFKSTSTTGLTSVSWSPQEGADTTVSFFHYSQGV 95  
QY 83 EVES---PFQGRLOWNGSKDLQDVSTVLNVLTNDLSGLYTCNVSRFEFEHRRPFVKTTTL 140  
Db 96 YLGNYPFKRISWAGDLKKDASINENMQFIHNGTYICDVKN----- 139  
QY 141 IPLRVTEAGEDFTSVVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQGHIRLYVVEKENLPVFPVWVVGIVTAVVLGTLTLLSMILAVLYRKNKR 198  
QY 188 E 188  
Db 199 D 199

## RESULT 10

US-09-430-503-18  
; Sequence 18, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Homo sapiens

## US-09-430-503-18

Query Match 11.3%; Score 127; DB 4; Length 209;  
Best Local Similarity 26.0%; Pred. No. 7.8e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;  
QY 27 VEVPSETEAVQGNPKRLCISCMKREVEAT---TVVEWVYRPEGGKDFL-IYEYRNGHQ 82  
Db 40 VYTPKEIFVANGTQGLKTC---KFKSTSTTGLTSVSWSPQEGADTTVSFFHYSQGV 95  
QY 83 EVES---PFQGRLOWNGSKDLQDVSTVLNVLTNDLSGLYTCNVSRFEFEHRRPFVKTTTL 140  
Db 96 YLGNYPFKRISWAGDLKKDASINENMQFIHNGTYICDVKN----- 139  
QY 141 IPLRVTEAGEDFTSVVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQGHIRLYVVEKENLPVFPVWVVGIVTAVVLGTLTLLSMILAVLYRKNKR 198  
QY 188 E 188  
Db 199 D 199

## RESULT 11

US-09-430-503-20  
; Sequence 20, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-20

Query Match 11.3%; Score 127; DB 4; Length 209;  
Best Local Similarity 26.0%; Pred. No. 7.8e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;  
QY 27 VEVPSETEAVQGNPKRLCISCMKREVEAT---TVVEWVYRPEGGKDFL-IYEYRNGHQ 82  
Db 40 VYTPKEIFVANGTQGLKTC---KFKSTSTTGLTSVSWSPQEGADTTVSFFHYSQGV 95  
QY 83 EVES---PFQGRLOWNGSKDLQDVSTVLNVLTNDLSGLYTCNVSRFEFEHRRPFVKTTTL 140  
Db 96 YLGNYPFKRISWAGDLKKDASINENMQFIHNGTYICDVKN----- 139  
QY 141 IPLRVTEAGEDFTSVVSE-----IMMYLLVFLTLMLLIEMIYCY---RKVSKA 187  
Db 140 -PPDIVVQGHIRLYVVEKENLPVFPVWVVGIVTAVVLGTLTLLSMILAVLYRKNKR 198  
QY 188 E 188  
Db 199 D 199

## RESULT 12

US-09-430-503-22  
; Sequence 22, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME  
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2  
; CURRENT APPLICATION NUMBER: US/09/430,503

; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-22

Query Match 11.3%; Score 127; DB 4; Length 209;  
Best Local Similarity 26.0%; Pred. No. 7.8e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;  
QY 27 VEVPSETEAVQGNPMKLRICISCKREVEAT---TVVEWFYRPEGKDFL-IYEYRNGHQ 82  
Db 40 VYTPKEIFVANGTQGLKTC---KFKSTSTGGLTSVMSFQEGADTTVSFFHYSGQGV 95  
QY 83 EVES--PFQGRLOWNGSKLDQVSIITVLNVTLDNSGLYTCNVSRFEFEAHRPFVKTRL 140  
Db 96 YLGNYPFKDRISWAGDLKDKDASINENMQFIHNGTYICDVKN-----139  
QY 141 IPLRVTEAGEDTSVVSE-----IMMVLIVFLTLMLLIEMICY---RKVSKA 187  
Db 140 -PPDIVVQGHIRLYVVEKENLPFPVWVVGIVTAVVLGLTLLISMLAVIYRKNKR 198  
QY 188 E 188  
Db 199 D 199

## RESULT 13

US-09-430-503-24  
; Sequence 24, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-430-503-24

Query Match 11.3%; Score 127; DB 4; Length 209;  
Best Local Similarity 26.0%; Pred. No. 7.8e-06;  
Matches 47; Conservative 21; Mismatches 73; Indels 40; Gaps 7;  
QY 27 VEVPSETEAVQGNPMKLRICISCKREVEAT---TVVEWFYRPEGKDFL-IYEYRNGHQ 82  
Db 40 VYTPKEIFVANGTQGLKTC---KFKSTSTGGLTSVMSFQEGADTTVSFFHYSGQGV 95  
QY 83 EVES--PFQGRLOWNGSKLDQVSIITVLNVTLDNSGLYTCNVSRFEFEAHRPFVKTRL 140  
Db 96 YLGNYPFKDRISWAGDLKDKDASINENMQFIHNGTYICDVKN-----139  
QY 141 IPLRVTEAGEDTSVVSE-----IMMVLIVFLTLMLLIEMICY---RKVSKA 187  
Db 140 -PPDIVVQGHIRLYVVEKENLPFPVWVVGIVTAVVLGLTLLISMLAVIYRKNKR 198  
QY 188 E 188  
Db 199 D 199

## RESULT 14

US-09-430-503-26  
; Sequence 26, Application US/09430503

; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-430-503-26

Query Match 11.1%; Score 125; DB 4; Length 270;  
Best Local Similarity 27.0%; Pred. No. 1.9e-05;  
Matches 47; Conservative 23; Mismatches 78; Indels 26; Gaps 8;  
QY 27 VEVPSETEAVQGNPMKLRICISCKREVEAT---TVVEWFYRPEGKDFL-IYEYRNGHQ 82  
Db 40 VHTPKKEIFVNGTQGLKTCF---DSPNTTGLTTSVMSFQEGDTSVAVSFHYSGQGV 95  
QY 83 EVES--PFQGRLOWNGSKLDQVSIITVLNVTLDNSGLYTCNVSRFEFEAHRP-----F 134  
Db 96 YIGDYPPFKDRVTWAGDLKDKDASINENIQAVHNGTYICDVKNPPDIVV-RPGHRLHV 154  
QY 135 VKTRLPLRVTEAGEDTSVVSEIMMYILLVFLTLMLLIEMICYRKVSKAE 188  
Db 155 VEIDNLLVFLVWVVG-TVTAVVLGLTLLISLVLVLY-----RRKHSKRD 199

## RESULT 15

US-09-430-503-28  
; Sequence 28, Application US/09430503  
; Patent No. 6355786  
; GENERAL INFORMATION:  
; APPLICANT: Zhao, Zhizhuang  
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND  
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME  
; CURRENT APPLICATION NUMBER: US/09/430,503  
; CURRENT FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 28  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-430-503-28

Query Match 11.1%; Score 125; DB 4; Length 270;  
Best Local Similarity 27.0%; Pred. No. 1.9e-05;  
Matches 47; Conservative 23; Mismatches 78; Indels 26; Gaps 8;  
QY 27 VEVPSETEAVQGNPMKLRICISCKREVEAT---TVVEWFYRPEGKDFL-IYEYRNGHQ 82  
Db 40 VHTPKKEIFVNGTQGLKTCF---DSPNTTGLTTSVMSFQEGDTSVAVSFHYSGQGV 95  
QY 83 EVES--PFQGRLOWNGSKLDQVSIITVLNVTLDNSGLYTCNVSRFEFEAHRP-----F 134  
Db 96 YIGDYPPFKDRVTWAGDLKDKDASINENIQAVHNGTYICDVKNPPDIVV-RPGHRLHV 154  
QY 135 VKTRLPLRVTEAGEDTSVVSEIMMYILLVFLTLMLLIEMICYRKVSKAE 188  
Db 155 VEIDNLLVFLVWVVG-TVTAVVLGLTLLISLVLVLY-----RRKHSKRD 199

Search completed: December 5, 2003, 23:24:06  
Job time : 31 secs



Db 121 NVSRFEFEAHRPFVKTTRILPLRVTEAGEDFTSVSEIMMYILLVFLTLMLLIEMIYC 180  
Qy 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

## RESULT 2

US-10-142-201B-11  
; Sequence 11, Application US/10142201B  
; Publication No. US20030022205A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals Inc.  
; TITLE OF INVENTION: A SODIUM CHANNEL BETA 4 SUBUNIT,  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MP12001-106P1RN(M)  
; CURRENT APPLICATION NUMBER: US/10/142,201B  
; CURRENT FILING DATE: 2002-05-09  
; PRIOR APPLICATION NUMBER: US 60/289,893  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-142-201B-11

Query Match 100.0%; Score 1124; DB 15; Length 215;  
Best Local Similarity 100.0%; Pred. No. 1.4e-109;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPAFNRLPLASLVLIYVWVCFVCPVCEVSETEAVQGNPKMLRCISCMKREVEATTVV 60  
Db 1 MPAFNRLPLASLVLIYVWVCFVCPVCEVSETEAVQGNPKMLRCISCMKREVEATTVV 60  
Qy 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITLVNLTNDGLTYC 120  
Db 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITLVNLTNDGLTYC 120  
Qy 121 NVSRFEFEAHRPFVKTTRILPLRVTEAGEDFTSVSEIMMYILLVFLTLMLLIEMIYC 180  
Db 121 NVSRFEFEAHRPFVKTTRILPLRVTEAGEDFTSVSEIMMYILLVFLTLMLLIEMIYC 180  
Qy 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
Db 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

## RESULT 3

US-09-997-579-1  
; Sequence 1, Application US/09997579  
; Patent No. US20020113203A1  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge University Technical Services  
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodium channel  
; TITLE OF INVENTION: channel  
; FILE REFERENCE: nucleic acids encoding them and therapeutic or diagnostic uses thereof  
; CURRENT APPLICATION NUMBER: US/09/997,579  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: PCT/EP00/01783  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60,129,473  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-997-579-1

Query Match 98.3%; Score 1105; DB 10; Length 215;  
Best Local Similarity 98.1%; Pred. No. 1.3e-107;  
Matches 211; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MPAFNRLPLASLVLIYVWVCFVCPVCEVSETEAVQGNPKMLRCISCMKREVEATTVV 60  
Db 1 MPAFNRLPLASLVLIYVWVCFVCPVCEVSETEAVQGNPKMLRCISCMKREVEATTVV 60  
Qy 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITLVNLTNDGLTYC 120  
Db 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITLVNLTNDGLTYC 120  
Qy 121 NVSRFEFEAHRPFVKTTRILPLRVTEAGEDFTSVSEIMMYILLVFLTLMLLIEMIYC 180  
Db 121 NVSRFEFEAHRPFVKTTRILPLRVTEAGEDFTSVSEIMMYILLVFLTLMLLIEMIYC 180  
Qy 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
Db 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

## RESULT 4

US-10-029-191-2  
; Sequence 2, Application US/10029191  
; Publication No. US20020160453A1  
; GENERAL INFORMATION:  
; APPLICANT: CURTIS, RORY A.J.  
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
; FILE REFERENCE: 210147.00XX/5U1  
; CURRENT APPLICATION NUMBER: US/10/029,191  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 09/569,978  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/134,198  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-029-191-2

Query Match 98.3%; Score 1105; DB 14; Length 215;  
Best Local Similarity 98.1%; Pred. No. 1.3e-107;  
Matches 211; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MPAFNRLPLASLVLIYVWVCFVCPVCEVSETEAVQGNPKMLRCISCMKREVEATTVV 60  
Db 1 MPAFNRLPLASLVLIYVWVCFVCPVCEVSETEAVQGNPKMLRCISCMKREVEATTVV 60  
Qy 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITLVNLTNDGLTYC 120  
Db 61 EMFYRPEGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQDVSIITLVNLTNDGLTYC 120  
Qy 121 NVSRFEFEAHRPFVKTTRILPLRVTEAGEDFTSVSEIMMYILLVFLTLMLLIEMIYC 180  
Db 121 NVSRFEFEAHRPFVKTTRILPLRVTEAGEDFTSVSEIMMYILLVFLTLMLLIEMIYC 180  
Qy 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215  
Db 181 YRKVSKAEAAQENASDYLAIPSENKENSAPVVEE 215

## RESULT 5

US-10-142-201B-12  
; Sequence 12, Application US/10142201B  
; Publication No. US20030022205A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals Inc.  
; APPLICANT: CURTIS, RORY A.J.

; TITLE OF INVENTION: 98359, A SODIUM CHANNEL BETA 4 SUBUNIT,  
; FILE REFERENCE: AND USES THEREFOR  
; FILE REFERENCE: MP12001-106P1RN(M)  
; CURRENT APPLICATION NUMBER: US/10/142,201B  
; CURRENT FILING DATE: 2002-05-09  
; PRIOR APPLICATION NUMBER: US 60/289,893  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-142-201B-12

Query Match 98.3%; Score 1105; DB 15; Length 215;  
Best Local Similarity 98.1%; Pred. No. 1.3e-107;  
Matches 211; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MPAPNRLPLASLVLIYVSVCFVCPVSETEAVQGNPKLRCISCMKREEVEATTVV 60  
DB 1 MPAPNRLPLASLVLIYVSVCFVCPVSETEAVQGNPKLRCISCMKREEVEATTVV 60  
QY 61 EWFYRPEGKDFLIYVRNGHQEVESPFQGRLOWNGSKDLQDVSTVNLVTLNDSGLYTC 120  
DB 61 EWFYRPEGKDFLIYVRNGHQEVESPFQGRLOWNGSKDLQDVSTVNLVTLNDSGLYTC 120  
QY 121 NVSRFEFEFAHRPFVKTTLRILPLRVTEEAGEDFTSVSEIMMYILLVFLTLMLFIEMIVC 180  
DB 121 NVSRFEFEFAHRPFVKTTLRILPLRVTEEAGEDFTSVSEIMMYILLVFLTLMLFIEMIVC 180  
QY 181 YKVKSAEAAQENASDYLAIPSEKNSAVPVEE 215  
DB 181 YKVKSAEAAQENASDYLAIPSEKNSAVPVEE 215

RESULT 6  
US-10-029-191-4  
; Sequence 4, Application US/10029191  
; Publication No. US20020160453A1  
; GENERAL INFORMATION:  
; APPLICANT: CURTIS, ROY A.J.  
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
; FILE REFERENCE: 210147.00XX/501  
; CURRENT APPLICATION NUMBER: US/10/029,191  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 09/569,978  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/134,198  
; PRIOR FILING DATE: 1993-05-14  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-029-191-4

Query Match 87.7%; Score 986; DB 14; Length 191;  
Best Local Similarity 99.0%; Pred. No. 3.4e-95;  
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 25 VCVVEPSETEAVQGNPKLRCISCMKREEVEATTVVWFYRPEGKDFLIYVRNGHQEV 84  
DB 1 VCVVEPSETEAVQGNPKLRCISCMKREEVEATTVVWFYRPEGKDFLIYVRNGHQEV 60  
QY 85 ESWFGRLOWNGSKDLQDVSTVNLVTLNDSGLYTCNVSRFEFEFAHRPFVKTTLRILPLR 144  
DB 61 ESWFGRLOWNGSKDLQDVSTVNLVTLNDSGLYTCNVSRFEFEFAHRPFVKTTLRILPLR 120  
QY 145 VTEEAGEDFTSVSEIMMYILLVFLTLMLFIEMIVCYRKYKSKAEAAQENASDYLAIPSE 204

DB 121 VTEEAGEDFTSVSEIMMYILLVFLTLMLFIEMIVCYRKYKSKAEAAQENASDYLAIPSE 180  
QY 205 NKENSAPVVEE 215  
DB 181 NKENSAPVVEE 191

## RESULT 7

US-09-997-579-22  
; Sequence 22, Application US/09997579  
; Patent No. US20020113203A1  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge University Technical Services  
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated so  
; TITLE OF INVENTION: channel  
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses  
; FILE REFERENCE: 674558-2001  
; CURRENT APPLICATION NUMBER: US/09/997,579  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: PCT/EP00/01783  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60,129,473  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 22  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-997-579-22

Query Match 75.0%; Score 843; DB 10; Length 159;  
Best Local Similarity 100.0%; Pred. No. 2.5e-80;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPAPNRLPLASLVLIYVSVCFVCPVSETEAVQGNPKLRCISCMKREEVEATTVV 60  
DB 1 MPAPNRLPLASLVLIYVSVCFVCPVSETEAVQGNPKLRCISCMKREEVEATTVV 60  
QY 61 EWFYRPEGKDFLIYVRNGHQEVESPFQGRLOWNGSKDLQDVSTVNLVTLNDSGLYTC 120  
DB 61 EWFYRPEGKDFLIYVRNGHQEVESPFQGRLOWNGSKDLQDVSTVNLVTLNDSGLYTC 120  
QY 121 NVSRFEFEFAHRPFVKTTLRILPLRVTEEAGEDFTSVVSE 159  
DB 121 NVSRFEFEFAHRPFVKTTLRILPLRVTEEAGEDFTSVVSE 159

## RESULT 8

US-09-997-579-23  
; Sequence 23, Application US/09997579  
; Patent No. US20020113203A1  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge University Technical Services  
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated so  
; TITLE OF INVENTION: channel  
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses  
; FILE REFERENCE: 674558-2001  
; CURRENT APPLICATION NUMBER: US/09/997,579  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: PCT/EP00/01783  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60,129,473  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 23  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-997-579-23

Query Match

74.0%; Score 832; DB 10; Length 159;

Best Local Similarity 98.7%; Pred. No. 3.6e-79;  
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1	MPAENRLEPLASLVLIYWSVCFPVCEVSETEAVQGNPKURCISCMKREBEVATVV	60
QY		
DB		
1	MPAFNRLLPLASLVLIYWRVCFPVCEVSETERVQGNPKURCISCMKREBEVATVV	60
QY		
DB		
61	EWFYRPEGGKFLIYEYRNGHQEVSEPFQRLQWNGSKLDQVSIIVLVNLTNDSGLYTC	120
QY		
DB		
61	EWFYRPEGGKFLIYEYRNGHQEVSEPFQRLQWNGSKLDQVSIIVLVNLTNDSGLYTC	120
QY		
DB		
121	NVSRSEFEFAHRPFVKLTRELPLRVTEAGEDFTSVVSE	159
QY		
DB		
121	NVSRSEFEFAHRPFVKLTRELPLRVTEAGEDFTSVVSE	159
QY		
DB		

```

RESULT 9
US-10-029-191-5
; Sequence 5, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, ROY A. J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE OF INVENTION: PROTEIN
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-029-191-5

```

Query Match	53.1%	Score 597;	DB 14;	Length 111;
Best Local Similarity	100.0%;	Pred. No. 9.2e-55;		
Matches 111;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

  

QY	25	VCVEVPSEAVCGNPMKLCISCMKREEVEATTVVEWFYRPEGGKDFLIYRYNGHQEV	84
Db	1	VCVEVPSEAVCGNPMKLCISCMKREEVEATTVVEWFYRPEGGKDFLIYRYNGHQEV	60
QY	85	ESPFGRLQWNGSKLDQYSITVTLVNTLDSGLYTCNVSREFFFEAHRPPV	135
Db	61	ESPFGRLQWNGSKLDQYSITVTLVNTLDSGLYTCNVSREFFFEAHRPPV	111

RESULT 10  
US-10-029-191-20  
; Sequence 20, Application US/10029191  
; Publication NO. US20020160453A1  
; GENERAL INFORMATION:  
; APPLICANT: CURTIS, ROY A.J.  
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 210147.00XX/SUI  
; CURRENT APPLICATION NUMBER: US/10/029,191  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 09/569,978  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/134,198  
; PRIOR FILING DATE: 1993-05-14  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Lepus Sp.

US-10-029-191-20

Query Match	42.4%;	Score 477;	DB 14;	Length 219;
Best Local Similarity	49.5%;	Pred. No. 8.8e-42;		
Matches 106;	Conservative 32;	Mismatches 68;	Indels 8;	Gaps 5;
Qy	10	LASLVLIYVYSCFFVCVEVPSETEAVQGNPKMLRCISCMKREVEATTVVEMFYRPEGG	69	
Db	5	LAPVGAALVSSANGCGVEVDSETEAVYGNWTFKILCISKRSSETTASFTSTFRQKGT	64	
Qy	70	KDPL-ITYVRNGHQEVESP--FGRLQWNGS-----KQLQDVSTTVLVNTLVNDSGLYTCNVS	123	
Db	65	EEFVKILRYENEVLQEEDERFEGRVVWNGSRGKQLQLSIPITNTVYNHSGDYQCHVY	124	
Qy	124	REFEFAHRFPVKTTRLPLRVTEBAGEDFTSVVSRIMMYVLAVFETLMLILMIYCYRK	183	
Db	125	RLLSFNYEHNTSVVKIHLHVVDKARNDWASIVSSIMMYVLIVLTILVLAEMVTCYKK	184	
Qy	184	VSKA-BEAAQENASDYLAIPSENKEN-SAVPVEE	215	
Db	185	IAAATEAAQAENASEYLAITSSKENCTGVQVAE	218	

```

RESULT 11
US-09-997-579-44
; Sequence 44, Application US/09997579
; Patent No. US20020113203A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit
; TITLE OF INVENTION: channel
; TITLE OF INVENTION: nucleic acids encoding them
; FILE REFERENCE: 674558-2001
; CURRENT APPLICATION NUMBER: US/09/997,579
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Rat
US-09-997-579-44

```

Query Match	42.0%;	Score 472;	DB 10;	Length 218;
Best Local Similarity	49.1%;	Pred. No. 2.9e-41;		
Matches 105; Conservative 33; Mismatches 68; Indels 8; Gaps 5;				
<hr/>				
QY	10	LASLVLIYVSVCFPVCEPSETAEVAQCNPMLRCISCMKREEEVATTVWENFYEPGG	69	
Db	5	LALVGC AVLSSANGCCEVDSETEAVYGWTFKILCISCRRSETTAETFTFWFRQGT	64	
QY	70	KDPL-IYEVRNGHQEVSP--FCGRLOWNGS---KDLDVSIITVLNVTLLNDGSLYTCNVYS	123	
		:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::
Db	65	EBFVKILRYENEVLQLEEDEREPEGRVWNNGSRGT KDQLDSIFITNTVYNHSGDYECVVY	124	
QY	124	REPEPAHRPFVKTTRFLIRLYTERAGGEDFTSVVSEIMMYILLVFLTLWLLEMIYCVRK	183	
Db	125	RLLFDNFDEHTSVVKKTHLEVVDKARDMASIVSEIMMYLIVLTTLVAEMVICYKK	184	
QY	184	VSKA-EAAQAENASDYLAIPSENKEN-SAVPVEE	215	
		:: :: :: :: ::	:: :: :: :: ::	:: :: :: :: ::
Db	185	IATAATEAAQAENASEYLAIITSKEKNCITGVQAAE	218	

RESULT 12  
US-10-142-201B-8  
; Sequence 8, Application US/10142201B  
; Publication No. US2003002205A1  
; GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals Inc.  
APPLICANT: Curtis, Rory A.J.  
TITLE OF INVENTION: 98359, A SODIUM CHANNEL BETA 4 SUBUNIT,  
TITLE OF INVENTION: AND USES THEREFOR  
FILE REFERENCE: MPI2001-106PRN(W)  
CURRENT APPLICATION NUMBER: US/10/142, 201B  
CURRENT FILING DATE: 2002-05-09  
PRIOR APPLICATION NUMBER: US 60/289, 893  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 218  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-142-201B-8

Query Match 41.9%; Score 471; DB 15; Length 218;  
Best Local Similarity 49.5%; Pred. No. 3.7e-41;  
Matches 106; Conservative 31; Mismatches 69; Indels 8; Gaps 5;  
Qy 10 LASLVLIYVWSVCFPCVVEPSETEAVQGNPKLCISCKRKEVEATTVVWFYRPRGG 69  
Db 5 LALVVGALVSSACGGCEVDETEAVYGMFKILCISCKRSEINASTFTFTWFRQGT 64  
Qy 70 KDFL-IYEYRNGHVESP--FQRLQWNGS---KDLQDVSTVTLNLTNDGLTYCNVS 123  
Db 65 BEFVKILRYENEVLQLEEDERPEGRVWNGSRGTKDLQDLSIFITVNTYHSGDYECHVY 124  
Qy 124 REFEFAHRPFVKTRLIPLRYTERAGDFTSVSEIMMYILLVLTLLIEMTYCYRK 183  
Db 125 RLLFFENYHNTSVVKKIHIEVVDKANRDMASIVSEIMMYILLVLTLLIEMTYCYK 184  
Qy 184 VSKA-EAAQENASDYLAIIPSENKEN-SAVPVEE 215  
Db 185 IAAATETAQENASEYLAITSKENCCTGVQVAE 218

RESULT 13  
US-10-029-191-7  
Sequence 7, Application US/10029191  
Publication No. US20020160453A1  
GENERAL INFORMATION:  
APPLICANT: CURTIS, RORY A.J.  
TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
TITLE OF INVENTION: PROTEIN  
FILE REFERENCE: 210147.00XX/SU1  
CURRENT APPLICATION NUMBER: US/10/029,191  
CURRENT FILING DATE: 2001-12-20  
PRIOR APPLICATION NUMBER: 09/569,978  
PRIOR FILING DATE: 2000-05-12  
PRIOR APPLICATION NUMBER: US 60/134,198  
PRIOR FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Rattus sp.  
US-10-029-191-7

Query Match 25.1%; Score 282; DB 14; Length 58;  
Best Local Similarity 96.6%; Pred. No. 3.8e-22;  
Matches 56; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 158 SEIMMYILLVLTLLIEMTYCYRKSKAEAAQENASDYLAIIPSENKENSAPVVEE 215  
Db 1 SEIMMYILLVLTLLIEMTYCYRKSKAEAAQENASDYLAIIPSENKENSAPVVEE 58

RESULT 14  
US-09-875-456A-14  
Sequence 14, Application US/09875456A

Patent No. US20020045229A1  
GENERAL INFORMATION:  
APPLICANT: Qin, Ning  
APPLICANT: Codd, Ellen  
APPLICANT: D'Andrea, Michael  
TITLE OF INVENTION: DNAs encoding human betala sodium channel subunit  
FILE REFERENCE: ORT-1221  
CURRENT APPLICATION NUMBER: US/09/875,456A  
CURRENT FILING DATE: 2001-09-10  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 268  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-875-456A-14

Query Match 23.0%; Score 258.5; DB 9; Length 268;  
Best Local Similarity 43.2%; Pred. No. 9.2e-19;  
Matches 64; Conservative 20; Mismatches 57; Indels 7; Gaps 4;  
Qy 10 LASLVLIYVWSVCFPCVVEPSETEAVQGNPKLCISCKRKEVEATTVVWFYRPRGG 69  
Db 5 LALVVGALVSSACGGCEVDETEAVYGMFKILCISCKRSEINASTFTFTWFRQGT 64  
Qy 70 KDFL-IYEYRNGHVESP--FQRLQWNGS---KDLQDVSTVTLNLTNDGLTYCNVS 123  
Db 65 BEFVKILRYENEVLQLEEDERPEGRVWNGSRGTKDLQDLSIFITVNTYHSGDYECHVY 124  
Qy 124 REFEFAHRPFVKTRLIPLRYTERAGE 151  
Db 125 RLLFFENYHNTSVVKKIHIEVVDK-GE 151

RESULT 15  
US-09-997-579-17  
Sequence 17, Application US/09997579  
Patent No. US20020113203A1  
GENERAL INFORMATION:  
APPLICANT: Cambridge University Technical Services  
TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated so  
TITLE OF INVENTION: channel  
TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses  
FILE REFERENCE: 674558-2001  
CURRENT APPLICATION NUMBER: US/09/997,579  
CURRENT FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: PCT/EP00/01783  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: 60,129,473  
PRIOR FILING DATE: 2000-02-24  
NUMBER OF SEQ ID NOS: 47  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 39  
TYPE: PRT  
ORGANISM: homo sapiens  
US-09-997-579-17

Query Match 18.4%; Score 207; DB 10; Length 39;  
Best Local Similarity 100.0%; Pred. No. 1.6e-14;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MPANRLLFPPLASLVLIYVWSVCFPCVVEPSETEAVQGN 39  
Db 1 MPANRLLFPPLASLVLIYVWSVCFPCVVEPSETEAVQGN 39

Search completed: December 5, 2003, 23:25:36  
Job time : 79 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:16:11 ; Search time 5264 Seconds  
(without alignments)  
9799.969 Million cell updates/sec

Title: US-09-936-680-4

Perfect score: 1261

Sequence: 1 cccctccctccgagctgagc.....tgccagaactgagaagccgg 1261

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pi.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pin.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1261	100.0	1261	6	AX039100	Sequence
2	1261	100.0	1261	9	HS243396	AX243396 Homo sapi
3	1252.2	99.3	4052	9	HS243396	AL136589 Homo sapi
4	909.4	72.1	5306	9	AB032984	AB032984 Homo sapi
5	835.8	66.3	3296	9	AB097521	AB097521 Macaca fa
6	559.6	44.4	2632	6	AX048005	AX048005 Sequence
7	559.6	44.4	3107	10	AF378093	AF378093 Rattus no
8	559.6	44.4	3108	6	AX047984	AX047984 Sequence
9	540.6	42.9	870	10	AY049036	AY049036 Mus muscu
10	540	42.8	2220	6	AX039099	AX039099 Sequence
11	540	42.8	2220	10	RN0243395	AJ243395 Rattus no
12	534.6	42.4	645	6	AX048004	AX048004 Sequence
13	429	34.0	471	6	BD059018	BD059018 Secreter
14	428	33.9	127347	2	AP000682	AP000682 Homo sapi
15	428	33.9	144833	2	AC063921	AC063921 Homo sapi
16	428	33.9	149800	2	AC021981	AC021981 Homo sapi
17	428	33.9	172546	2	AC024604	AC024604 Homo sapi
18	428	33.9	178169	9	AP002765	AP002765 Homo sapi
19	428	33.9	181471	9	AC069539	AC069539 Homo sapi
20	405	32.1	178431	2	AP002749	AP002749 Homo sapi
21	229.6	18.2	172546	2	AC024604	AC024604 Homo sapi
22	228	18.1	144833	2	AC063921	AC063921 Homo sapi
23	228	18.1	149800	2	AC021981	AC021981 Homo sapi
24	216	17.1	178431	2	AP002749	AP002749 Homo sapi
25	179	14.2	143624	2	AC135353	AC135353 Mus muscu
26	177	14.0	66980	2	AC013796	AC013796 Homo sapi
27	174.2	13.8	214306	2	AC128723	AC128723 Rattus no
28	148.2	11.8	657	4	OC035382	U35382 Oryctolagus
29	148.2	11.8	657	6	AX048006	AX048006 Sequence
30	145.4	11.5	1404	9	HUMVSCN1BA	L10338 Human sodiu
31	143.8	11.4	1335	9	HUMVSCN1B	L16242 Homo sapien
32	138.6	11.0	1125	10	MMU46681	U46681 Mus musculu
33	138.6	11.0	1340	10	MMU85786	U85786 Mus musculu
34	138.6	11.0	1490	6	AX01978	AX01978 Sequence
35	138.6	11.0	1490	10	RATSCN1B	BC039140 Mus muscu
36	138.6	11.0	1557	10	BC039140	BC039140 Mus muscu
37	138.6	11.0	1568	10	BC009652	BC009652 Mus muscu
38	129	10.2	66980	2	AC013796	AC013796 Homo sapi
39	119.6	9.5	1248	9	BC021266	BC021266 Homo sapi
40	92.6	7.3	74467	2	AC013674	AC013674 Homo sapi
41	86.4	6.9	143624	2	AC135353	AC135353 Mus muscu
42	81.4	6.5	807	6	AX354520	AX354520 Sequence
43	81.4	6.5	974	6	AX354521	AX354521 Sequence
44	74.4	5.9	136892	5	AL954724	AL954724 Zebrafish
45	73	5.8	850	6	AX105630	AX105630 Sequence

ALIGNMENTS

RESULT 1	AX039100	Sequence 4 from Patent WO0063367.	1261 bp	DNA	linear	PAT 18-NOV-2000
LOCUS	AX039100	Sequence 4 from Patent WO0063367.				
DEFINITION	AX039100					
ACCESSION	AX039100					
VERSION	AX039100.1	GI:11229276				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Cox, P.; Dixon, A.; Jackson, A. and Morgan, K.					
TITLE	A novel family of beta sub-unit proteins from a voltage-gated sodi					
	um channel, nucleic acids encoding them and therapeutic or					

diagnostic uses there of  
 Patent: WO 0063367-A 4 26-OCT-2000;  
 WARNER-LAMBERT COMPANY (US); Cambridge University Technical  
 Services Limited (GB)  
 FEATURES Location/Qualifiers  
 source 1..1261  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 263 a 366 c 372 g 260 t  
 ORIGIN  
 Query Match 100.0%; Score 1261; DB 6; Length 1261;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-288;  
 Matches 1261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CCCTCCCTTCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGGCGCGAGTGG 60  
 DB 1 CCCTCCCTTCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGGCGCGAGTGG 60  
 QY 61 AAGCTGGAGTTCCGGGTTGGCGGAGCGGAGCTGTCGTGTCGAGCGCGCGAGGA 120  
 DB 61 AAGCTGGAGTTCCGGGTTGGCGGAGCGGAGCTGTCGTGTCGAGCGCGCGAGGA 120  
 QY 121 GCGGCGCGGAGCGGAGCTGAGCTGCTCCCTCGAACTGGGAGGTCAGTGGGTCGCTTAG 180  
 DB 121 GCGGCGCGGAGCGGAGCTGAGCTGCTCCCTCGAACTGGGAGGTCAGTGGGTCGCTTAG 180  
 QY 181 GGCCCAAGCCCCCAGCCGCTCCCAAGCTCCAGAGCCCTCCCGAGCGACCGGTCTCG 240  
 DB 181 GGCCCAAGCCCCCAGCCGCTCCCAAGCTCCCGAGCGCTCCCGAGCGACCGGTCTCG 240  
 QY 241 GCCCTCTCTCGTCCAGAGCTCGCCCTCGGCGGAGTTCGTCGCAAGGTTCTCTCG 300  
 DB 241 GCCCTCTCTCGTCCAGAGCTCGCCCTCGGCGGAGTTCGTCGCAAGGTTCTCTCG 300  
 QY 301 AAAGAACTCTGAGAGCGGCGAGCTCTTGACCGAGGGAATCTCTGTGTAGCTTTGGAAGC 360  
 DB 301 AAAGAACTCTGAGAGCGGCGAGCTCTTGACCGAGGGAATCTCTGTGTAGCTTTGGAAGC 360  
 QY 361 GCGCAGCCCCAGAGATCCCTGCTTCAATAGATGTTTCCCTGCTGCTCTCTGCTGCTT 420  
 DB 361 GCGCAGCCCCAGAGATCCCTGCTTCAATAGATGTTTCCCTGCTGCTCTCTGCTGCTT 420  
 QY 421 ATCTACTGGTTCAGTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 DB 421 ATCTACTGGTTCAGTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 481 GTGAGGCGCAACCCCATGAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 DB 481 GTGAGGCGCAACCCCATGAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 QY 541 GCCACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 DB 541 GCCACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 601 GAGTATCGGAATGGCCACGAGGAGTGGAGAGCGCTTTCAGGGCGGCTGCTGCTGCTGCTG 660  
 DB 601 GAGTATCGGAATGGCCACGAGGAGTGGAGAGCGCTTTCAGGGCGGCTGCTGCTGCTGCTG 660  
 QY 661 GCGCAGAGGACCTCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 DB 661 GCGCAGAGGACCTCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 QY 721 GGCCTCTACCTCGAATGTCGCGGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAG 780  
 DB 721 GGCCTCTACCTCGAATGTCGCGGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAG 780  
 QY 781 AAGACGACCGGCTGATCCCCCTAAGAGTCAACCGAGGAGGCTGGAGAGGACTTCACTCT 840  
 DB 781 AAGACGACCGGCTGATCCCCCTAAGAGTCAACCGAGGAGGCTGGAGAGGACTTCACTCT 840  
 QY 841 GTGGTCTCAGAAATCATGATGATACATCTCTCTGCTCTCTCAACCTCTGCTGCTCATC 900

841 GTGGTCTCAGAAATCATGATGATACATCTCTGCTCTCTCAACCTCTGCTGCTCATC 900  
 QY 901 GAGATGATATATTGCTTACAGAAAGGTCTCAAAAGCCGGAAGAGGAGCCCAAGAAACGCG 960  
 DB 901 GAGATGATATATTGCTTACAGAAAGGTCTCAAAAGCCGGAAGAGGAGCCCAAGAAACGCG 960  
 QY 961 TCTGACTACTTGGCCATCCCATCTGACACAGGAGGAGTCTGCGGTACAGTGGAGGAA 1020  
 DB 961 TCTGACTACTTGGCCATCCCATCTGACACAGGAGGAGTCTGCGGTACAGTGGAGGAA 1020  
 QY 1021 TAGACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTTGAGGGAGTGGACATCCCATG 1080  
 DB 1021 TAGACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTTGAGGGAGTGGACATCCCATG 1080  
 QY 1081 TTGAGCAATGTCAATGGCATCAGAGGCGGCCCAAGGGGCCCATCGCTTCCCTTATGTC 1140  
 DB 1081 TTGAGCAATGTCAATGGCATCAGAGGCGGCCCAAGGGGCCCATCGCTTCCCTTATGTC 1140  
 QY 1141 ATCCATGTTCTGTTTCAATTCATCCATACATCCACCTGCTCTGAGCTTTCACTCT 1200  
 DB 1141 ATCCATGTTCTGTTTCAATTCATCCATACATCCACCTGCTCTGAGCTTTCACTCT 1200  
 QY 1201 GACTCCCTAACTCAGACCTCTACGACCAATAGACTCTGCCAGAACTTGAGAGCGC 1260  
 DB 1201 GACTCCCTAACTCAGACCTCTACGACCAATAGACTCTGCCAGAACTTGAGAGCGC 1260  
 QY 1261 G 1261  
 DB 1261 G 1261  
 RESULT 2  
 HSA243396  
 LOCUS Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit  
 DEFINITION (scn3b gene).  
 ACCESSION AJ243396  
 VERSION 2 GI-7242612  
 KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K.,  
 Finnock, R.D., Hines, J., Richardson, P.J., Mizuguchi, K. and  
 Jackson, A.P.  
 beta 3: an additional auxiliary subunit of the voltage-sensitive  
 sodium channel that modulates channel gating with distinct kinetics  
 Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)  
 20160948  
 PUBMED 10698874  
 REFERENCE 2 Morgan, K.  
 Direct Submission  
 Submitted (28-JUN-1999) Morgan K., Biochemistry, University of  
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
 Revised by [4]  
 3 (bases 1 to 1261)  
 REFERENCE Morgan, K.  
 Direct Submission  
 Submitted (13-MAR-2000) Morgan K., Biochemistry, University of  
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
 On Mar 14, 2000 this sequence version replaced gi:7160974.  
 COMMENT Location/Qualifiers  
 FEATURES 1..1261  
 source /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="brain"  
 1..1261  
 /gene="scn3b"  
 gene





```

/dev_stage="adult"
1..5306
/gene="KIAA1158"
<1..695
/gene="KIAA1158"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1158 protein"
/protein_id="BAA86472.1"
/db_xref="GI:6330136"
/translation="RSMGSGSLGRQKMPAFNRLPLASLVLYVSVCPVCEV
PSEIARVQNMKRCISCMREVEATVWFYRPEGKDFLIYRNGHQBESV
FQRLQNGKLDQVSIIVLNLNDISGLYCNVSEFEFPAHFVKTTRLPLRV
TEEAGEFTSVSVSIMYILVFLTLMLLIEMICYRKVKAEEAQAENASDYLAI
ENKENSAPVPEE"
BASE COUNT 1375 a 1311 c 1111 g 1509 t
ORIGIN

Query Match 72.1%; Score 909.4; DB 9; Length 5306;
Best Local Similarity 99.9%; Pred. No. 9.7e-205;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 349 AGCCTTGAAAGCGCCAGCCAGAGATGCTGCTTCAATAGATTGTTCCCTCGCT 408
DB 21 AGCCTTGAAAGCGCCAGCCAGAGATGCTGCTTCAATAGATTGTTCCCTCGCT 80
QY 409 TCTCTCGTGTATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTCG 468
DB 81 TCTCTCGTGTATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTCG 140
QY 469 GAGACGAGGCGGTGACGGCAACCCATGAGTGGCTGCTCATCTCTGTCATGAAGA 528
DB 141 GAGACGAGGCGGTGACGGCAACCCATGAGTGGCTGCTCATCTCTGTCATGAAGA 200
QY 529 GAGGAGGTGGAGGCGCCAGCAGCTGTGGATGTGTTACAGGCGCGAGGGCGGTAAAGT 588
DB 201 GAGGAGGTGGAGGCGCCAGCAGCTGTGGATGTGTTACAGGCGCGAGGGCGGTAAAGT 260
QY 589 TTCCTTATTTACGAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGC 648
DB 261 TTCCTTATTTACGAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGC 320
QY 649 CTGACAGTGAATGACGAGGAGCTGTCAGAGCTGTCTCATCTGCTCAAGTCACT 708
DB 321 CTGACAGTGAATGACGAGGAGCTGTCAGAGCTGTCTCATCTGCTCAAGTCACT 380
QY 709 CTGAACGACTCTGCGCTCTACACCTGCAATGTGTCGGGAGTTTGAGTTTGAGGCGCAT 768
DB 381 CTGAACGACTCTGCGCTCTACACCTGCAATGTGTCGGGAGTTTGAGTTTGAGGCGCAT 440
QY 769 CGGCGCTTTGTGAAGACGACGCGCTGATCCCTTAAGAGTCAACGAGGAGGCTGGAGAG 828
DB 441 CGGCGCTTTGTGAAGACGACGCGCTGATCCCTTAAGAGTCAACGAGGAGGCTGGAGAG 500
QY 829 GACTTCACTCTGTGGTCTGAGAAATCATGATGATGATCTCTGCTTCTTCTTCCCTG 888
DB 501 GACTTCACTCTGTGGTCTGAGAAATCATGATGATGATCTCTGCTTCTTCTTCCCTG 560
QY 889 TGGCTGCTCATCGAGATGATATTTGCTACAGAAAGTCTCAAAAGCGGAGAGGAGGCC 948
DB 561 TGGCTGCTCATCGAGATGATATTTGCTACAGAAAGTCTCAAAAGCGGAGAGGAGGCC 620
QY 949 CAAGAAAACGGTCTGACTTCCATCCATCCATCTGAGAACAGAGAACTCTGCGGTA 1008
DB 621 CAAGAAAACGGTCTGACTTCCATCCATCCATCTGAGAACAGAGAACTCTGCGGTA 680
QY 1009 CCAGTGGAGGAATAGAACAGGAGCAGTGTGACATGAGTGGCTGCTGACACCTGAGGACT 1068
DB 681 CCAGTGGAGGAATAGAACAGGAGCAGTGTGACATGAGTGGCTGCTGACACCTGAGGACT 740
QY 1069 GGACATCCCATGTTACCAATGTCAATGGCATCAGGAGGCGGCCCAAGGGCCCCCATCGC 1128
DB 741 GGACATCCCATGTTACCAATGTCAATGGCATCAGGAGGCGGCCCAAGGGCCCCCATCGC 800

QY 1129 TTCCCTTCATGCATCCATTTGTTCTGTTTCATTCATTCATCATCATCATCCACTGCTCTGA 1188
DB 801 TTCCCTTCATGCATCCATTTGTTCTGTTTCATTCATTCATCATCATCATCCACTGCTCTGA 860
QY 1189 GCTTTCACCTCTGACTCCCTAACTCCATCAGAGCTCTACGACCATTAAGACTCTGCCAGA 1248
DB 861 GCTTTCACCTCTGACTCCCTAACTCCATCAGAGCTCTACGACCATTAAGACTCTGCCAGA 920
QY 1249 ACTGAGAGGCC 1259
DB 921 ACTGAGAGGCC 931

RESULT 5
AB097521
LOCUS
DEFINITION
Macaca fascicularis brain cDNA clone:QmoA-13657, full insert
sequence.
ACCESSION
AB097521
VERSION
AB097521.1 GI:26449236
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Macaca fascicularis (crab-eating macaque)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirata,M.,
Suto,Y., Hirai,M., Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
chromosomes
JOURNAL
Gene 275 (1), 31-37 (2001)
MEDLINE
21458551
PUBMED
11574149
REFERENCE
2 (bases 1 to 3296)
Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
Direct Submission
Submitted (05-DEC-2002) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACCATGTG)
R. Site2: DraIII (CACCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by oligo-capping method. Custom primers used for
sequencing ( 5' end primer [CTTCTGCTTAAAGGCTGG],
3' end primer [CGACCTGCAGTCGACACA] ).
FEATURES
Location/Qualifiers
1..3296
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="QmoA-13657"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="macaque brain cDNA library QmoA"
/dev_stage="adult"
6..653
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC41746.1"
/db_xref="GI:26449237"
/translation="MPAFNRLPLVSLVLIYWSVCFVCEVFPSETBAVQGNPMKLR"

CDS
```

C1SCMKREBEVATTVVWFYRREGGKDFLIYVRNGHQBVESFPQRLQWNGSKLOLD  
VSTLVNVLNDGLYCNVSEFEFAHRPFVKTRLLPLRVTEAGDFTSVVSEI  
MMVILLVFLTLMLLIEMIIYKRVKSAEBAQENASDYLAIPISENKENSAPVEE"  
BASE COUNT 886 a 830 c 699 g 881 t  
ORIGIN

Query Match 66.3%; Score 835.8; DB 9; Length 3296;  
Best Local Similarity 97.1%; Pred. No. 2.7e-187;  
Matches 863; Conservative 0; Mismatches 22; Indels 4; Gaps 1;  
QY 371 AGAAGATGCCCTGCTCAATAGATGTTTCCCTGGCTCTCTCGTCTATCTACTGG 430  
DB 1 AGAAGATGCCCTGCTCAACAGATGTTTCCCTGGCTCTCTCGTCTATCTACTGG 60  
QY 431 TCAGTGTCTGCTTCCCTGTGTGTGAAGTGCCTTCGAGAGCGAGCGGCGGCA 490  
DB 61 CCAGTGTCTGCTTCCCTGTGTGTGAAGTGCCTTCGAGAGCGAGCGGCTGTG 120  
QY 491 ACCCCATGAAGCTGCGCTCATCTCTGATGAAGAGAGAGAGGAGGAGGAGGAG 550  
DB 121 ACCCCATGAAGCTGCGCTCATCTCTGATGAAGAGAGAGAGGAGGAGGAGGAG 180  
QY 551 TGTGTGAATGGTGTCTACAGCCCGAGGCGGTAAAGATTTCCTATTACGATATCGA 610  
DB 181 TGTGTGAATGGTGTCTACAGCCCGAGGCGGTAAAGATTTCCTATTACGATATCGA 240  
QY 611 ATGGCCACAGAGGTTGAGAGGCGCTTTCAGGGGCGCTGAGTGAATGGAGCAAG 670  
DB 241 ATGGCCACAGAGGTTGAGAGGCGCTTTCAGGGGCGCTGAGTGAATGGAGCAAG 300  
QY 671 ACCTGCAGAGACGTGTCCATCACTGTCTCAAGCTCACTCTGAACGACTCTGGCCCTACA 730  
DB 301 ACCTGCAGAGGTTGAGAGGCGCTTTCAGGGGCGCTGAGTGAATGGAGCAAG 360  
QY 731 CTGCAATGTGTCCTGGGAGTTGAGTTTGAAGCGCATCGGCCCTTTGTGAAGACGAC 790  
DB 361 CTGCAATGTGTCCTGGGAGTTGAGTTTGAAGCGCATCGGCCCTTTGTGAAGACGAC 420  
QY 791 GCGTATCCCTTAAGAGTCAACGAGGAGGCTGAGAGGACTTCACTCTGTGGTCTAG 850  
DB 421 GCGTATCCCTTAAGAGTCAACGAGGAGGCTGAGAGGACTTCACTCTGTGGTCTAG 480  
QY 851 AAATCATGATGTATCATCTCTGCTTCTCACCCTGTGGCTGTCTCATCGAGATGATAT 910  
DB 481 AAATCATGATGTATCATCTCTGCTTCTCACCCTGTGGCTGTCTCATCGAGATGATAT 540  
QY 911 ATGCTACAGAAAGTCTCAAAAGCGGAGAGCGGCCCAAGAAAGCGGTCTGACTACC 970  
DB 541 ATGCTACAGAAAGTCTCAAAAGCGGAGAGCGGCCCAAGAAAGCGGTCTGACTACC 600  
QY 971 TTGGCATCCCATCTGAGAACAGGAGAACTCTGCGGTACCTGAGGAGGAGGAGGAG 1030  
DB 601 TTGGCATCCCATCTGAGAACAGGAGAACTCTGAGTACCTGAGGAGGAGGAGGAG 660  
QY 1031 GAGTGTGATGATGAGTGGCTGAAACACTGAGGAGGAGTGGACATCCCATGTTTCA 1090  
DB 661 GAGTGTGATGATGAGTGGCTGAAACACTGAGGAGGAGTGGACATCCCATGTTTCA 720  
QY 1091 TCAATGGATCAGAGGAGGCGCCCAAGGCGCCCATCGCTTCCCTTCATGATCCATGTT 1150  
DB 721 TCAATGGATCAGAGGAGGCGCCCAAGGCGCCCATCGCTTCCCTTCATGATCCATGTT 780  
QY 1151 CTGTTTCATTCATTCATTCATATATCCACCTGCTCTGAGCTTTTCACTCTGACTCCCTAA 1210  
DB 781 CT----ATTTCATTCATTCATATATCCACCTGCTCTGAGCTTTTCACTCTGACTCCCTAA 836  
QY 1211 CTCCATCAGACCTCTACGACCATTAAGACTCTGCGAGAACTGAGAGGCC 1259  
DB 837 CTCCATCAGACCTCTACGACCATTAAGACTCTGCGAGAACTGAGAGGCC 885

RESULT 6  
AX048005

LOCUS AX048005 2632 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 22 from Patent WO0069912.  
ACCESSION AX048005  
VERSION AX048005.1 GI:11876883  
KEYWORDS Rattus sp.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1  
AUTHORS Curtis, R.A.  
TITLE Gene encoding a sodium channel beta-3 subunit protein  
JOURNAL Patent: WO 0069912-A 22 23-NOV-2000;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..2632  
/organism="Rattus sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10118"  
BASE COUNT 693 a 636 c 611 g 692 t  
ORIGIN

Query Match 44.4%; Score 559.6; DB 6; Length 2632;  
Best Local Similarity 84.4%; Pred. No. 7.8e-122;  
Matches 655; Conservative 0; Mismatches 114; Indels 7; Gaps 2;  
QY 371 AGAAGATGCCCTGCTCAATAGATGTTTCCCTGGCTCTCTCGTCTATCTACTGG 430  
DB 73 AAAGATGCCCTGCTTCAACAGATGTTTCCCTAGCTTCTTAGTGTCTATCTACTGG 132  
QY 431 TCAGTGTCTGCTTCCCTGTGTGTGAAGTGCCTTCGAGACGAGGCGGTCGAGGCA 490  
DB 133 TCAGAGTCTGCTTCCCTGTGTGTGAAGTGCCTTCGAGACGAGGCGGTCGAGGCA 192  
QY 491 ACCCATCAAGCTGCGCTCATCTCTGATGAAGAGAGAGGAGGAGGAGGAGGAGG 550  
DB 193 ATCCCATGAAGCTGAGGTGATCTCTGATGAAGAGGAGGAGGAGGAGGAGGAGG 252  
QY 551 TGTGTGAATGTTTCTACAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 610  
DB 253 TGTGTGAGTGTCTTACAGGCGCTGAGGCGGTTAAGATTTCTTATATAGTATCGGA 312  
QY 611 ATGCGCACAGAGGTTGAGAGGCGCTTTTCAAGGCGCGCTGCGAGTGAATGCGAGCAAG 670  
DB 313 ATGCGCACAGAGGTTGAGAGGCGCTTTTCAAGGCGCGCTGCGAGTGAATGCGAGCAAG 372  
QY 671 ACCTGCAGAGTGTCTCATCACTGTCTCAAGCTCACTCTGAACGAGCTCTGCGCTCTACA 730  
DB 373 ACCTGCAGAGTGTCTCATCACTGTCTCAATGTCATCTTGAATGACTCTGCGCTCTACA 432  
QY 731 CCTCAATGTGTCCGGAGTTTGAAGGCGCATCGGCCCTTTTGTGAAGACGAGC 790  
DB 433 CATCAATGTGTCCAGGAGTTTGAATTCGAGGCGCACACAGGCGCTTTTGTGAAGACGAGC 492  
QY 791 GCGTATCCCTTAAGAGTCAACGAGGCGCTGAGAGGAGTTCACCTCTGTGTCTCAG 850  
DB 493 GACTGATACCTTTCCGAGTCACTGAGAGGCGGAGAGGAGTTCACCTCTGTGTCTCGG 552  
QY 851 AAATCATGATGATATCTCTCTGTGTCTTCTCACCCTGTGTGTGTGTGTGTGTGTGTGT 910  
DB 553 AAATCATGATGATATCTCTCTGTGTCTTCTCACCCTGTGTGTGTGTGTGTGTGTGTGT 612  
QY 911 ATTGCTACAGAAAGGTTCTCAAAAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 970  
DB 613 ATTGCTACAGAAAGGTTCTTAAAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 672  
QY 971 TTGCGATCCCATCTGAGAACAGAGGAGGAGTCTGCGGTACCGAGTGGAGGAGGAGGAGGAG 1030  
DB 673 TTGCTATCCCTTTCAGAGAACAGAGGAGGAGTCTGTGGTACCTGTGGTGGAGGAGGAGG 726  
QY 1031 GCAGTGTGACATGAGGTGGGCTTGAACACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1090







Db 478 ATCCATGAGCTGAGTGTGCTCTCTGCTATGAGAGGAGGAGTGGAGGCCACCACTG 537  
QY 551 TGTGTGAATGGTCTACAGGCCCGAGCGGTAAGATTTCTTTATTTACGAGTATCGGA 610  
Db 538 TGTGTGAATGGTCTACAGGCCCGAGCGGTAAGATTTCTTTATTTACGAGTATCGGA 597  
QY 611 ATGCCACACAGGAGTGGAGAGCCCTTTTCAGGGGCGCTGCGAGTGGAGTGGAGCAGG 670  
Db 598 ATGCCACACAGGAGTGGAGAGCCCTTTTCAGGGGCGCTGCGAGTGGAGTGGAGCAGG 657  
QY 671 ACTGCGAGGAGTGTCCATCACTGTGCTCAACGCTCACTCTGAACGACTCTGGGCTCTACA 730  
Db 658 ACTGCGAGGAGTGTCCATCACTGTGCTCAACGCTCACTCTGAACGACTCTGGGCTCTACA 717  
QY 731 CCTGCAATGTGTCGCGGAGTTTGTGTTGAGCGCATGCGCCCTTTGTGAAGACGACGC 790  
Db 718 CATGCAATGTGTCGCGGAGTTTGTGTTGAGCGCATGCGCCCTTTGTGAAGACGACGC 777  
QY 791 GGCTGATCCCTTAAGAGTCAACGAGGAGCTGGAGGAGTTCACCTCTGTGCTCTCAG 850  
Db 778 GACTGATACCTTTGCGAGTCACTGAAGAGCGGAGAGACTTCACCTCTGTGCTCTCAG 837  
QY 851 AATATGATGATGATCATCTCTCTGCTCTCTCAACCTGTGCTGCTCATGAGATGATAT 910  
Db 838 AATATGATGATGATCATCTCTCTGCTCTCTCAACCTGTGCTGCTCATGAGATGATAT 897  
QY 911 ATTGCTACGAGAGGTTCTCAAGAGCGGAGGAGCGCCCAAGAGGAGTTCGACTACC 970  
Db 898 ATTGCTACGAGAGGTTCTCAAGAGCGGAGGAGCGCCCAAGAGGAGTTCGACTACC 957  
QY 971 TTCCATCCATCTGAGAACAGGAGAGTCTGCGGTACCACTGCTGAGGAGTAT 1022  
Db 958 TTGCTATCCCTCAGAGAACAGGAGAGTCTGCTGAGTACCTGTGAGGAGTAT 1009

RESULT 12  
AX048004  
LOCUS AX048004 645 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 21 from Patent WO0069912.  
ACCESSION AX048004  
VERSION AX048004.1 GI:11876882  
KEYWORDS Lepus sp.  
SOURCE Lepus sp.  
ORGANISM Lepus sp.  
REFERENCE 1  
AUTHORS Curtis,R.A.  
TITLE Gene encoding a sodium channel beta-3 subunit protein  
JOURNAL Patent: WO 0069912-A 21 23-NOV-2000;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..645  
/organism="Lepus sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:137772"  
BASE COUNT 155 a 154 c 181 g 155 t  
ORIGIN  
Query Match 42.4%; Score 534.6; DB 6; Length 645;  
Best Local Similarity 89.3%; Pred. No. 6.6e-116;  
Matches 576; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 376 ATGCGTGCCTTCAATAGATGTTTCCCTGGCTTCTCTGCTGCTTATCTACTGGGTCACT 435  
Db 1 ATGCGTGCCTTCAATAGATGTTTCCCTGGCTTCTCTGCTGCTTATCTACTGGGTCACT 60  
QY 436 GTCTGCTTCCCTGTGTGTGAAGTGCCTCGAGAGCGGCGGTGAGGCAACCC 495  
Db 61 GTCTGCTTCCCTGTGTGTGAAGTGCCTCGAGAGCGGCGGTGAGGCAACCC 120  
QY 496 ATGAGGCTGCGTGTGATCTCTCTGATGAGAGAGGAGGTGAGGCGCCACCACTGGTGTG 555

Db 121 ATGAAGCTGAGTGTGATCTCTCTGCTATGAGAGGAGGAGTGGAGGCCACCACTGTG 180  
QY 556 GAATGGTCTTACAGGCCCGAGCGGTAAGATTTCTTTATTTACGAGTATCGGAATGC 615  
Db 181 GAGTGGTCTTACAGGCCCGAGCGGTAAGATTTCTTTATTTACGAGTATCGGAATGC 240  
QY 616 CACCGAGGAGTGGAGAGCCCTTTTCAGGGGCGCTGCGAGTGGAGTGGAGCAGCCTG 675  
Db 241 CACCGAGGAGTGGAGAGCCCTTTTCAGGGGCGCTGCGAGTGGAGTGGAGCAGCCTG 300  
QY 676 CAGGAGTGTCCATCACTGTGCTCAACGCTCACTCTGAACGACTCTGGGCTCTTACACCTG 735  
Db 301 CAGGAGTGTCCATCACTGTGCTCAACGCTCACTCTGAACGACTCTGGGCTCTTACACCTG 360  
QY 736 AATGTGTCCCGGAGTGTGAGTTTGAAGCGCATGCGCCCTTTGTGAAGAGCAGCGGCTG 795  
Db 361 AATGTGTCCCGGAGTGTGAGTTTGAAGCGCATGCGCCCTTTGTGAAGAGCAGCGGCTG 420  
QY 796 ATCCCGCTTAAGAGTCAACGAGGAGCTGGAGAGGCTTCACCTCTGTGCTCTCAGAAATC 855  
Db 421 ATACCTTTGCGAGTCACTGAAGAGCGGAGAGACTTCACCTCGTGTGCTCTCGGAAATC 480  
QY 856 ATGATGTACATCTCTCTGCTCTCTCAACCTGTGCTGCTCATCGAGTATGATATATTC 915  
Db 481 ATGATGTACATCTCTCTGCTCTCTCAACCTGTGCTGCTCATCGAGTATGATATATTC 540  
QY 916 TACGAGAGGTTCTCAAGAGCGGAGAGGAGCGCCCAAGAGAGGCTGTGACTACCTTGC 975  
Db 541 TACGAGAGGTTCTCTAAGCGCGGAGAGGAGCGCCCAAGAGAGGCTGTGACTACCTTGC 600  
QY 976 ATCCCATCTGAGAGACAGGAGACTCTCGGTACCACTGAGTGGAGGAA 1020  
Db 601 ATCCCTTTCAGAGAACAGGAGACTCTGTGTACTCTGTGAGGAA 645

RESULT 13  
BD059018  
LOCUS BD059018 471 bp DNA linear PAT 27-AUG-2002  
DEFINITION Secreted expressed sequence tags (seSTs).  
ACCESSION BD059018  
VERSION BD059018.1 GI:22604624  
KEYWORDS Zea mays  
SOURCE Zea mays  
ORGANISM Zea mays  
REFERENCE 1  
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D., Tracy,M., Spaulding,V. and Agostino,M.J.  
TITLE Secreted expressed sequence tags (seSTs)  
JOURNAL Patent: JP 2001519666-A 873 23-OCT-2001;  
GENETICS INSTITUTE INC  
COMMENT PN JP 2001519666-A/873  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI DAVID MERBERG,  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
1..471  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
BASE COUNT 93 a 126 c 143 g 108 t 1 others  
ORIGIN  
Query Match 34.0%; Score 429; DB 6; Length 471;

```

Best Local Similarity 99.5%; Pred. No. 7.1e-91;
Matches 440; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 369 CCAGAGATGCTGCTTCAATAGATTTTCCCTGGCTTCTCTCGTCTTACTG 428
Db 22 CAGAGATGCTGCTTCAATAGATTTTCCCTGGCTTCTCTCGTCTTACTG 81

QY 429 GGTGATGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGCCGTCAGGG 488
Db 82 GGTGATGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGG-GCGTGCAGGG 140

QY 489 CACCCCATGAGAGTGGCTGCTGCTTCAATAGATTTTCCCTGGCTTCTCTCGTCTTACTG 548
Db 141 CACCCCATGAGAGTGGCTGCTGCTTCAATAGATTTTCCCTGGCTTCTCTCGTCTTACTG 200

QY 549 GGTGATGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGCCGTCAGGG 608
Db 201 GGTGATGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGG-GCGTGCAGGG 260

QY 609 GAATGCTGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGCCGTCAGGG 668
Db 261 GAATGCTGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGG-GCGTGCAGGG 320

QY 669 GACCTGCTGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGCCGTCAGGG 728
Db 321 GACCTGCTGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGG-GCGTGCAGGG 380

QY 729 CACCTGCTGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGCCGTCAGGG 788
Db 381 CACCTGCTGCTGCTTCCCTGCTGTTGTTGAGTGCCTTCGAGACGGG-GCGTGCAGGG 440

QY 789 GCGGCTGATCCCCCTAAGAGTC 810
Db 441 GCGGCTGATCCCCCTAAGAGTC 462

RESULT 14
AP000682/c
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone CMB9-32A1 map 11q24, WORKING DRAFT
SEQUENCE 8 unordered pieces.
ACCESSION
AP000682.3 GI:9844967
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 127347)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 127,347 genomic DNA of 11q24
Published Only in DataBase (1999)
2 (bases 1 to 127347)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (08-NOV-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail: hattori@gsr.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-42-778-9923, Fax: 81-42-778-9924)
On Aug 18, 2000 this sequence version replaced gi:8118870.
COMMENT
-----
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsr.riken.go.jp
-----
Project Information
Center project name: HumDraft11
Center clone name: CMB9-32A1
-----
Sequencing vector: PCR products; 100% of reads

```

Chemistry: Dye-terminator ET-amersham: 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 124981 bases at least Q40  
 Consensus quality: 125815 bases at least Q30  
 Consensus quality: 126286 bases at least Q20  
 Insert size: 126647; sum-of-contigs  
 Quality coverage: 12.71x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 33532 contig of 33532 bp in length  
 33633 59223 contig of 25591 bp in length  
 59324 33633 59223 contig of 23684 bp in length  
 83108 100196 contig of 17089 bp in length  
 100297 112313 contig of 12017 bp in length  
 112414 121043 contig of 8630 bp in length  
 121444 125917 contig of 4774 bp in length  
 126018 127347 contig of 1330 bp in length.

\* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 33532: contig of 33532 bp in length  
 33633: gap of 100 bp  
 59223: contig of 25591 bp in length  
 59324: gap of 100 bp  
 83108: contig of 23684 bp in length  
 83107: gap of 100 bp  
 100196: contig of 17089 bp in length  
 100197: gap of 100 bp  
 112313: contig of 12017 bp in length  
 112414: gap of 100 bp  
 121043: contig of 8630 bp in length  
 121444: gap of 100 bp  
 125917: contig of 4774 bp in length  
 126018: gap of 100 bp  
 127347: contig of 1330 bp in length.

Location/Qualifiers  
 1. 127347  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11Q24"  
 /clone="CMB9-32A1"  
 1. 33532  
 /note="assembly\_fragment"  
 33633. 59223  
 /note="assembly\_fragment"  
 59324. 83107  
 /note="assembly\_fragment"  
 83108. 100196  
 /note="assembly\_fragment clone\_end:SP6 vector\_side:left"  
 100297. 112313  
 /note="assembly\_fragment"  
 112414. 121043  
 /note="assembly\_fragment clone\_end:T7 vector\_side:right"  
 121144. 125917  
 /note="assembly\_fragment"  
 126018. 127347  
 /note="assembly\_fragment"  
 36325 a 27508 c 27825 g 34989 t 700 others  
 BASE COUNT  
 ORIGIN

Query Match 33.9% Score 428; DB 2; Length 127347;  
 Best local Similarity 100.0%; Pred. No. 1.2e-90;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTCCTTCGAGTGTAGCTTACCTGGGCGCAAAACGAGCGAGCGGCGGAGTGGAA 62  
 DB 29132 CTCCTTCGAGTGTAGCTTACCTGGGCGCAAAACGAGCGAGCGGCGGAGTGGAA 29073

QY 63 GCTGGAGTTCGCGGGTGGGCGGAGCGGAGTTCGCTGCTGTAGCGCGGCGAGAGC 122  
 DB 29072 GCTGGAGTTCGCGGGTGGGCGGAGCGGAGTTCGCTGCTGTAGCGCGGCGAGAGC 29013

QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAATCGGAGAGTTCAGTGGGTTCGCTAGGG 182  
 DB 29012 GGGCGGAGCGGCTGATCGGCTCCCTCGAATCGGAGAGTTCAGTGGGTTCGCTAGGG 28953

QY 183 CCAAGAGCCCGCCACCGGCTCCAAAGCTCCCAAGGCTCCCGAGGCGCTCCCGAGCGGCTCCGCG 242  
 DB 28952 CCAAGAGCCCGCCACCGGCTCCAAAGCTCCCGAGGCTCCCGAGGCGCTCCCGAGCGGCTCCGCG 28893

QY 243 CTTCTTCGTCAGAAAGTCGCGCCCTGGGCGAGTTCGTCCTCCAAAGGTTTCCTCGAA 302  
 DB 28892 CTTCTTCGTCAGAAAGTCGCGCCCTGGGCGAGTTCGTCCTCCAAAGGTTTCCTCGAA 28833

QY 303 AGAATCTGAGGCGGCGAGTTCGTCGAGGAGGAACTCTCTGTAGCGCTTGGAGCGG 362  
 DB 28832 AGAATCTGAGGCGGCGAGTTCGTCGAGGAGGAACTCTCTGTAGCGCTTGGAGCGG 28773

QY 363 CCAGCCCGCAGAGATGCTGCTCTCAATAGATGTTTCCCTGCTCTCTCTGCTGTTAT 422  
 DB 28772 CCAGCCCGCAGAGATGCTGCTCTCAATAGATGTTTCCCTGCTCTCTCTGCTGTTAT 28713

QY 423 CTACTGGG 430  
 DB 28712 CTACTGGG 28705

RESULT 15  
 AC063921/c  
 LOCUS AC063921 144833 bp DNA linear HTG 08-JAN-2003  
 DEFINITION Homo sapiens chromosome 3 clone RP11-142P10, WORKING DRAFT  
 SEQUENCE, 13 unordered pieces.  
 AC063921  
 VERSION AC063921.18 GI:20335587  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 144833)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
 Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Avele, M., Banks, T.,  
 Barbarella, J., Benton, J., Blum, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, A.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karisson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C.,  
 Kratovic, J., Kresh, A., Landry, N., Leal, B., Lewis, D.C., Lewis, L.,  
 Li, J., Li, Z., Lichtharge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,

Maheeswari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunolu, G.,  
 Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,  
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabox, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Umami, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 144833)  
 Worley, K.C.  
 Direct Submission  
 Submitted (22-APR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 144833)  
 Worley, K.C.  
 Direct Submission  
 Submitted (08-JAN-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Apr 28, 2002 this sequence version replaced gi:16117924.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: HAXY  
 Center clone name: RP11-142P10  
 ----- Summary Statistics  
 Sequencing vector: M13;  
 Chemistry: Dye-primer Body: 14% of reads  
 Chemistry: Dye-terminator Big Dye: 86% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 127260 bases at least Q40  
 Consensus quality: 132735 bases at least Q30  
 Consensus quality: 136183 bases at least Q20  
 Estimated insert size: 139078; sum-of-contigs estimation  
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 13 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 2437: contig of 2437 bp in length  
 \* 2438 2537: gap of unknown length  
 \* 2538 4812: contig of 2275 bp in length  
 \* 4813 4912: gap of unknown length  
 \* 4913 8488: contig of 3576 bp in length  
 \* 8489 8588: gap of unknown length  
 \* 8589 11163: contig of 2575 bp in length  
 \* 11164 11263: gap of unknown length  
 \* 11264 16638: contig of 5375 bp in length  
 \* 16639 16739: gap of unknown length  
 \* 16739 21870: contig of 5132 bp in length  
 \* 21871 21970: gap of unknown length

\* 21971 30780: contig of 8810 bp in length  
\* 30781 30880: gap of unknown length  
\* 30881 39267: contig of 8387 bp in length  
\* 39268 39367: gap of unknown length  
\* 39368 46890: contig of 7523 bp in length  
\* 46891 46990: gap of unknown length  
\* 46991 59656: contig of 12666 bp in length  
\* 59657 59756: gap of unknown length  
\* 59757 83109: contig of 23553 bp in length  
\* 83110 83409: gap of unknown length  
\* 83410 111857: contig of 28448 bp in length  
\* 111858 111957: gap of unknown length  
\* 111958 144833: contig of 32876 bp in length.

## FEATURES

## source

1..144833  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-142P10"

BASE COUNT 38456 a 33920 c 34115 g 37093 t 1249 others  
ORIGIN

Query Match 33.9%; Score 428; DB 2: Length 144833;

Best Local Similarity 100.0%; Pred. No. 1.2e-90;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CTCCTTCGAGCTGAGTTACCTGGCGCAACGAGCGAGGCGGCGCGAGTGGAA 62  
Db 50353 CTCCTTCGAGCTGAGTTACCTGGCGCAACGAGCGAGGCGGCGCGAGTGGAA 50294  
Qy 63 GCTGGAGTTCGGGGTGGCGGGGAGCGGACTGTCGGTGTCTGAGCGCGCGGAGAGC 122  
Db 50293 GCTGGAGTTCGGGGTGGCGGGGAGCGGACTGTCGGTGTCTGAGCGCGCGGAGAGC 50234  
Qy 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAGGG 182  
Db 50233 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAGGG 50174  
Qy 183 CCCAAGCCCCCACCCTCGGCTCCAAAGCTCCAGGGCTCCCGAGGACCGGTGCTCGGC 242  
Db 50173 CCCAAGCCCCCACCCTCGGCTCCAAAGCTCCAGGGCTCCCGAGGACCGGTGCTCGGC 50114  
Qy 243 CCTTCCTTCGGTCAGAAAGTCGCCCTCGGGGCGAGTTCGTCCCAAGGGTTTCCTCGAA 302  
Db 50113 CCTTCCTTCGGTCAGAAAGTCGCCCTCGGGGCGAGTTCGTCCCAAGGGTTTCCTCGAA 50054  
Qy 303 AGAATCTGAGAGGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCCTTGGAGCCG 362  
Db 50053 AGAATCTGAGAGGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCCTTGGAGCCG 49994  
Qy 363 CCAGCCCCAGAGATGCTGCTTCAATAGATTGTTCCCTGGCTTCTCTCGTCTTAT 422  
Db 49993 CCAGCCCCAGAGATGCTGCTTCAATAGATTGTTCCCTGGCTTCTCTCGTCTTAT 49934  
Qy 423 CTACTGGG 430  
Db 49933 CTACTGGG 49926

Search completed: December 5, 2003, 20:50:53  
Job time : 5274.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:12:31 ; Search time 475 Seconds  
(without alignments)  
7166.298 Million cell updates/sec

Title: US-09-936-680-4  
Sequence: 1 cctcctcccgagtgagc.....tgcagaactgagaagccgg 1261

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1261	100.0	1261	21 AAC67837	Human betas3 cDNA.
2	1252.2	99.3	4052	24 ABA93727	Human signal trans
3	909.4	72.1	1510	22 AAF84146	Human novel sodium
4	871.8	69.1	1045	22 AAK52345	Human polynucleoti
5	858.6	68.1	978	22 AAH98320	Human EST-derived
6	561.8	44.6	1195	23 AAS86764	DNA encoding novel
7	559.6	44.4	2632	22 AAC90602	Rat sodium channel
8	559.6	44.4	3108	22 AAC90600	Rat sodium channel

9	540	42.8	2220	21 AAC67836	Rat betas3 subunit
10	534.6	42.4	645	22 AAC90601	Rat sodium channel
11	429	34.0	471	20 AAY86895	EST clone BMA. Ho
12	392.6	31.1	3531	23 AAS86763	DNA encoding novel
13	224.4	17.8	621	23 AAS86762	DNA encoding novel
14	216	17.1	4625	22 AAL04971	Human reproductive
15	216	17.1	4625	23 ABL97864	Human testicular a
16	148.2	11.8	657	22 AAC90603	Rabbit sodium chan
17	142.2	11.3	1414	25 ABZ23837	Human voltage-gate
18	141	11.2	407	22 ABA08942	Human polynucleoti
19	141	11.2	407	22 AAK53329	Human polynucleoti
20	138.6	11.0	1490	24 ABK63747	Rat sequence diffe
21	81.4	6.5	974	24 AAD29622	Human betas3 sodiu
22	74	5.9	509	22 AAI93395	Human polynucleoti
23	73	5.8	850	22 AAF57675	Rat sodium channel
24	73	5.8	850	22 AAF30669	Sodium channel bet
25	45.4	3.6	1645	21 AAB88951	Maize geranylgeran
26	42.2	3.3	500	21 AAC41468	Zea mays DNA fragm
27	40.2	3.2	2806	20 AAZ40538	Full length human
28	40.2	3.2	2811	23 AAS83495	DNA encoding novel
29	39.8	3.1	2685	24 ABK37073	Bovine EST associa
30	39.6	3.1	2685	24 ABK3742	DNA encoding human
31	39.2	3.1	5973	24 ABS78752	Novel human cDNA s
32	39.2	3.1	6202	25 ABX70893	Human secretory po
33	39.2	3.1	7584	25 ABZ36402	Human secretory po
34	39	3.1	936	24 ABK73545	Bagillus lichenifo
35	39	3.1	3272	24 AB199475	Mouse ischaemic co
36	38.8	3.1	1299	23 ABL10323	Drosophila melanog
37	38.8	3.1	3787	23 ABL10322	Drosophila melanog
38	38.6	3.1	1179	23 AAS4061	Pseudomonas aerugi
39	38.6	3.1	1337	20 AAZ17263	Human gene express
40	38.6	3.1	1776	22 AAD06422	Human p21-activate
41	38.6	3.1	1941	21 AAZ40657	Human serine/threo
42	38.6	3.1	2310	25 ABX70915	Novel human cDNA s
43	38.6	3.1	2423	23 ABL24082	Drosophila melanog
44	38.6	3.1	4166	23 ABL24084	Drosophila melanog
45	38.6	3.1	11089	22 AAK79653	Human immune/haema

#### ALIGNMENTS

RESULT 1  
AAC67837  
ID AAC67837 standard; cDNA; 1261 BP.  
XX AAC67837;  
AC AAC67837;  
DT 15-FEB-2001 (first entry)  
DE Human betas3 cDNA.  
XX  
XX Human; beta sub-unit; betas3; analgesic; anticonvulsant;  
cerebroprotective; vasotropic; cardiant; nootropic; cytosstatic;  
dermatological; gene therapy; voltage-gated sodium channel; pain;  
epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;  
familial nonchromaffin paraganglioma; phenylketonuria;  
Charcot Marie tooth disease; ss.  
XX Homo sapiens.  
XX  
XX WQ200063367-A1.  
XX  
XX 26-OCT-2000.  
XX  
XX 24-FEB-2000; 2000MO-EP01783.  
XX  
XX 15-APR-1999; 99US-0129473.  
XX  
XX (WARN ) WARNER LAMBERT CO.  
XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.  
XX Cox P, Dixon A, Jackson A, Morgan K;  
PI

XX WPI; 2000-665241/64.  
 DR P-PSDB; AAB36002.  
 XX  
 PT Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
 PT sodium channel, and their corresponding polypeptides, useful for  
 PT detecting and treating sodium channel-associated conditions, e.g. pain,  
 PT epilepsy and stroke -  
 XX  
 PS Claim 10; Page 70-71; 88pp; English.  
 XX  
 CC The present sequence is given in the claims of a specification  
 CC relating to a novel family of beta sub-unit proteins from a  
 CC voltage-gated sodium channel. Human and rat beta sub-units, which  
 CC have been collectively identified as beta3, have been isolated.  
 CC The polynucleotides and polypeptides are useful for screening for  
 CC agonists and antagonists of sodium channels. The agonists, antagonists,  
 CC proteins and nucleic acids may be used diagnosing of treating diseases  
 CC or conditions associated with voltage-gated sodium channels, e.g. pain,  
 CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial  
 CC Nonchromaffin paraganglioma, Phenylketonuria and Charcot Marie Tooth  
 CC disease.  
 XX  
 SQ Sequence 1261 5P; 263 A; 366 C; 372 G; 260 T; 0 other;

Query Match 100.0%; Score 1261; DB 21; Length 1261;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTCCCTCCGAGTGGCTTACCTGGGGGCAACGAGCGAGGAGGCGGCGGAGTGG 60  
 DB 1 CCTCTCCCTCCGAGTGGCTTACCTGGGGGCAACGAGCGAGGAGGCGGCGGAGTGG 60  
 QY 61 AAGCTGAGATTCGCGGGTGGCGGGAGGCGAGCTGCTCCGTGGTCTGAGCGCGGAGGA 120  
 DB 61 AAGCTGAGATTCGCGGGTGGCGGGAGGCGAGCTGCTCCGTGGTCTGAGCGCGGAGGA 120  
 QY 121 GCGGGCGGAGGCGGCTGATCGGCTCCCTCGAAGTGGGAGGTCAGTGGGGTCTGTTAG 180  
 DB 121 GCGGGCGGAGGCGGCTGATCGGCTCCCTCGAAGTGGGAGGTCAGTGGGGTCTGTTAG 180  
 QY 181 GGGCCAAAGCCGCGGCTCCAAAGCTCCAGGGGCTCCGAGGCGGCGGAGGCTGCTG 240  
 DB 181 GGGCCAAAGCCGCGGCTCCAAAGCTCCAGGGGCTCCGAGGCGGCGGAGGCTGCTG 240  
 QY 241 GCGCTTCTCTCGGTGAGAAAGTCCGCTCCGAGGCTCCGAGGCGGCGGAGGCTGCTG 300  
 DB 241 GCGCTTCTCTCGGTGAGAAAGTCCGCTCCGAGGCTCCGAGGCGGCGGAGGCTGCTG 300  
 QY 301 AAAGAAATCTGAGAGGCGGAGTCTTGAACGAGGGAATCTCTCTGTGTAGCTTGAAGC 360  
 DB 301 AAAGAAATCTGAGAGGCGGAGTCTTGAACGAGGGAATCTCTCTGTGTAGCTTGAAGC 360  
 QY 361 GCGCAGGCGGAGGAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGT 420  
 DB 361 GCGCAGGCGGAGGAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGT 420  
 QY 421 ATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAAGTGGCTCCGAGAGCGAGGCG 480  
 DB 421 ATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAAGTGGCTCCGAGAGCGAGGCG 480  
 QY 481 GTGCAAGGCGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 DB 481 GTGCAAGGCGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 QY 541 GCGCAGGCGGAGGAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGT 600  
 DB 541 GCGCAGGCGGAGGAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGT 600  
 QY 601 GAGTATCGGAATGCGGAGGAGTGGAGAGGCGGCTTTCAGGGGCGGCTTTCAGGGGCG 660  
 DB 601 GAGTATCGGAATGCGGAGGAGTGGAGAGGCGGCTTTCAGGGGCGGCTTTCAGGGGCG 660

QY 661 GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 DB 661 GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 QY 721 GGCCTCTACACCTGCAATGTGTCCCGGAGTGTGAGTTTGAAGCGGAGGAGGAGGAGG 780  
 DB 721 GGCCTCTACACCTGCAATGTGTCCCGGAGTGTGAGTTTGAAGCGGAGGAGGAGGAGG 780  
 QY 781 AAGAGCAGCGGCTGATCCCTTAAGAGTCAACGAGGAGGAGGAGGAGGAGGAGGAGG 840  
 DB 781 AAGAGCAGCGGCTGATCCCTTAAGAGTCAACGAGGAGGAGGAGGAGGAGGAGGAGG 840  
 QY 841 GTGCTCTCAGAAATCATGATGTATCATCTTCTGCTTCTTCTTCTTCTTCTTCTTCTT 900  
 DB 841 GTGCTCTCAGAAATCATGATGTATCATCTTCTGCTTCTTCTTCTTCTTCTTCTTCTT 900  
 QY 901 GAGTATATATTTCTACAGAAAGTCTCAAAAGCGGAGGAGGAGGAGGAGGAGGAGG 960  
 DB 901 GAGTATATATTTCTACAGAAAGTCTCAAAAGCGGAGGAGGAGGAGGAGGAGGAGG 960  
 QY 961 TCTGACTACCTTGGCATCCCATCTGAGAACCAAGAGGAGGAGGAGGAGGAGGAGGAA 1020  
 DB 961 TCTGACTACCTTGGCATCCCATCTGAGAACCAAGAGGAGGAGGAGGAGGAGGAGGAA 1020  
 QY 1021 TAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080  
 DB 1021 TAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080  
 QY 1081 TTCAAGCAATGTCAATGGCATCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140  
 DB 1081 TTCAAGCAATGTCAATGGCATCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
 QY 1141 ATCAATTTCTGCTGCTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1200  
 DB 1141 ATCAATTTCTGCTGCTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1200  
 QY 1201 GACTCCCTTAACCTCAGACCTCTAGCGACCATTAAGACTCTGCGAGAACTGAGAGCGG 1260  
 DB 1201 GACTCCCTTAACCTCAGACCTCTAGCGACCATTAAGACTCTGCGAGAACTGAGAGCGG 1260  
 QY 1261 G 1261  
 DB 1261 G 1261

RESULT 2  
 ABA93727  
 ID ABA93727 standard; cDNA; 4052 BP.  
 XX  
 AC ABA93727;  
 XX  
 DT 30-APR-2002 (first entry)  
 XX  
 DE Human signal transduction cDNA clone amv2\_2f18.  
 XX  
 KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;  
 XX Gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200198454-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 25-APR-2001; 2001WO-IB02050.  
 XX  
 PR 25-APR-2000; 2000US-199380P.  
 XX  
 PA (GEHU-) GERMAN HUMAN GENOME PROJECT.  
 XX  
 PI Wiemann S;  
 XX  
 DR WPI; 2002-055860/07.

DR P-PSDB; ABB05689.  
 XX Human cDNA sequences and clones derived from human fetal brain, fetal  
 PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic  
 PT screening and therapy -  
 XX Claim 1; Page 174-175; 611pp; English.  
 XX The present invention describes assemblages and computer readable media  
 CC comprising novel human cDNA sequences and clones derived from human  
 CC fetal brain, fetal kidney, melanoma, testis and amygdala cDNA  
 CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the  
 CC present invention which encode the proteins given in ABB05662 to  
 CC ABB05729. The human cDNA sequences and clones can be used in gene  
 CC therapy. The clones may be used in a variety of applications, for  
 CC example they may be used in profiling assays, for providing large arrays  
 CC of human genetic material for implementing large-scale screening  
 CC strategies and for treating diseases via gene therapy procedures.  
 XX SQ Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 other;  
 Query Match 99.3%; Score 1252.2; DB 24; Length 4052;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1254; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 CTCCCTCCGAGCTGAGCTTACCTGGGCGCAACAGCGAGCGGCGGCGAGTGGAA 62  
 DB 431 CTCCTCCGAGCTGAGCTTACCTGGGCGCAACAGCGAGCGGCGGCGAGTGGAA 490  
 QY 63 GCTGAGTTCGGGGTGGGGGAGGAGGAGTCTGCTGGTGTGAGCGCGCGGAGAGC 122  
 DB 491 GCTGAGTTCGGGGTGGGGGAGGAGGAGTCTGCTGGTGTGAGCGCGCGGAGAGC 550  
 QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGGTTCGCTTAGGG 182  
 DB 551 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGGTTCGCTTAGGG 610  
 QY 183 CCCAAGCCCCACCCGCTCCAAAGCTCCAGGCGCTCCCGAGGCACTGGTGTGCGC 242  
 DB 611 CCCAAGCCCCCGCGCTCCAAAGCTCCAGGCGCTCCCGAGGCACTGGTGTGCGC 670  
 QY 243 CTTCTTCGCTCAGAAAGTCCGCGCGGCGAGTTCCTCCCAAGGGTTCCTCGAA 302  
 DB 671 CTTCTTCGCTCAGAAAGTCCGCGCGGCGAGTTCCTCCCAAGGGTTCCTCGAA 730  
 QY 303 AGAATCTCAGAGGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCCG 362  
 DB 731 AGAATCTCAGAGGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCCG 790  
 QY 363 CAGGCGCCAGAGATGCTGCTCCATAGATTCCTTCCCTGCTTCTCTGCTTAT 422  
 DB 791 CAGGCGCCAGAGATGCTGCTCCATAGATTCCTTCCCTGCTTCTCTGCTTAT 850  
 QY 423 CTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGAAAGTCCCTCGAGAGCGAGCGGT 482  
 DB 851 CTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGAAAGTCCCTCGAGAGCGAGCGGT 910  
 QY 483 CGAGGCGAAGCCCATGAGCTGCGCTGATCTCTGATGAGAGAGAGAGAGTGGAGGC 542  
 DB 911 CGAGGCGAAGCCCATGAGCTGCGCTGATCTCTGATGAGAGAGAGAGAGTGGAGGC 970  
 QY 543 CACACGCTGTGTGAATGGTCTACAGGCGCGAGGCGGTAAAGATTCCTATTATACGA 602  
 DB 971 CACACGCTGTGTGAATGGTCTACAGGCGCGAGGCGGTAAAGATTCCTATTATACGA 1030  
 QY 603 GTATCGAATGECACCAAGAGGAGTGGAGAGCCCTTTTCAGGGCGCTTCAGTGGATGG 662  
 DB 1031 GTATCGAATGECACCAAGAGGAGTGGAGAGCCCTTTTCAGGGCGCTTCAGTGGATGG 1090  
 QY 663 CAGCAGGAGCTGAGGAGGAGTCCATCAGTGTCTCAGTGTCTGAGAGAGCTGG 722  
 DB 1091 CAGCAGGAGCTGAGGAGGAGTCCATCAGTGTCTCAGTGTCTGAGAGAGCTGG 1150

QY 723 CCTCTACACCTGCAATGTGTCCTCCGGAGTTTGAGTTTGGAGCGCATCGCCCTTTGGAA 782  
 DB 1151 CCTCTACACCTGCAATGTGTCCTCCGGAGTTTGAGTTTGGAGCGCATCGCCCTTTGGAA 1210  
 QY 783 GACGACGGGCTGATTCCTCCCTAAGAGTCAACGAGGAGGCTGGAGAGACTTCACCTCTGT 842  
 DB 1211 GACGACGGGCTGATTCCTCCCTAAGAGTCAACGAGGAGGCTGGAGAGACTTCACCTCTGT 1270  
 QY 843 GGTCTCAGAAATCATGATGATCATCTTCTGCTTCTTCTCCTCCTTGGCTGCTCATCGA 902  
 DB 1271 GGTCTCAGAAATCATGATGATCATCTTCTGCTTCTTCTCCTCCTTGGCTGCTCATCGA 1330  
 QY 903 GATGATATATTGCTACAGAAAGGCTCTCAAAAGCCGAAAGAGGAGCGCCCAAGAAACGGGTC 962  
 DB 1331 GATGATATATTGCTACAGAAAGGCTCTCAAAAGCCGAAAGAGGAGCGCCCAAGAAACGGGTC 1390  
 QY 963 TGACTACTTGGCCATTCCTCATCTGAGAACAGAGAGACTCTGCGGTACAGTGGAGGAATA 1022  
 DB 1391 TGACTACTTGGCCATTCCTCATCTGAGAACAGAGAGACTCTGCGGTACAGTGGAGGAATA 1450  
 QY 1023 GAACAGGAGGAGTGTGACATGAGTGGCTCTGAACACCTGAGGGAGCTGGACATCCCATGTT 1082  
 DB 1451 GAACAGGAGGAGTGTGACATGAGTGGCTCTGAACACCTGAGGGAGCTGGACATCCCATGTT 1510  
 QY 1083 CAGCAATGTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1142  
 DB 1511 CAGCAATGTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1570  
 QY 1143 CCAATGTTCTGTTCAATGTTCAATGTTCAATGTTCAATGTTCAATGTTCAATGTTCAATGTT 1202  
 DB 1571 CCAATGTTCTGTTCAATGTTCAATGTTCAATGTTCAATGTTCAATGTTCAATGTTCAATGTT 1630  
 QY 1203 CTCCTTAATCTCATCAGACCTCTACGACCAATAGACTCTGCCAGAACTGAGAAGCC 1259  
 DB 1631 CTCCTTAATCTCATCAGACCTCTACGACCAATAGACTCTGCCAGAACTGAGAAGCC 1687  
 RESULT 3  
 AAF84146  
 ID AAF84146 standard; cDNA; 1510 BP.  
 XX AAF84146;  
 AC AAF84146;  
 XX DT 07-SEP-2001 (first entry)  
 DE Human novel sodium channel betal-like subunit encoding cDNA.  
 DE Sodium channel; sensory neurone specific channel; betal-like subunit;  
 KW SNS; therapeutic; pain; analgesic; ss.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 213..860  
 FT /\*tag= a  
 FT /product= "sodium channel betal-like subunit"  
 XX WO200144293-A2.  
 XX PD 21-JUN-2001.  
 XX 14-DEC-2000; 2000WO-GB04802.  
 XX 17-DEC-1999; 99GB-0029970.  
 XX (GLAX ) GLAXO GROUP LTD.  
 PA Plumpton M, Powell AJ, Sanseau P;  
 PI WPI: 2001-398129/42.  
 DR P-PSDB; AAB85206.  
 XX Novel sub-unit for voltage-gated sodium channel proteins for producing



Query Match 69.1%; Score 871.8; DB 22; Length 1045;  
Best Local Similarity 99.2%; Pred. No. 2.8e-224;  
Matches 876; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 86 GAGGCGACGTCTCGTGGTCTGAGCGCGCGGCGAGAGCGCGCGCGGCGGCTGATCGGCT 145  
DB 12 GTGGCGACGTCTCGTGGTCTGAGCGCGCGGCGAGAGCTGGCGCGGCGGCTGATCGGCT 71

QY 146 CCCTCGAAGTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGAGCCCGCCCGGCTCCA 205  
DB 72 CCCTCGAAGTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGAGCCCGCCCGGCTCCA 131

QY 206 AAAGCTCCAGAGGCTCCCGAGGACCGGTGCTCGGCCCTTCTTCGGTCAAGAGTGGC 265  
DB 132 AAAGCTCCAGAGGCTCCCGAGGACCGGTGCTCGGCCCTTCTTCGGTCAAGAGTGGC 191

QY 266 CCCTCGGAGGATGTCGTCGCGGAGGTTCTCGAAGATCTCGAGAGGCGGCGGCTGCT 325  
DB 192 CCCTCGGAGGATGTCGTCGCGGAGGTTCTCGAAGATCTCGAGAGGCGGCGGCTGCT 251

QY 326 TGACCGAGGAGTCTCTCTGTGTAGCTTGGAGCGCGCGAGCCCGAGAGATGCTGCT 385  
DB 252 TGACCGAGGAGTCTCTCTGTGTAGCTTGGAGCGCGCGAGCCCGAGAGATGCTGCT 311

QY 386 TCAATAGATGTTTCCCTCGGCTTCTCTGCTTATCTACTGCTGATGCTGCTTCC 445  
DB 312 TCAATAGATGTTTCCCTCGGCTTCTCTGCTTATCTACTGCTGATGCTGCTTCC 371

QY 446 CTGCTGTGAGGAGTCCCTCGAGAGCGGCGGCTGAGAGCGGCGGCGGCTGAGAGTGC 505  
DB 372 CTGCTGTGAGGAGTCCCTCGAGAGCGGCGGCTGAGAGCGGCGGCGGCTGAGAGTGC 431

QY 506 GCTGCATCTCTCTCATGAGAGAGGAGTGGAGGCGGCGGCGGCGGCTGAGAGTGC 565  
DB 432 GCTGCATCTCTCTCATGAGAGAGGAGTGGAGGCGGCGGCGGCGGCTGAGAGTGC 491

QY 566 ACAGGCGGAGGCGGCTGAGATTTCTTATTAAGTATGATGATGATGATGATGATGATG 625  
DB 492 ACAGGCGGAGGCGGCTGAGATTTCTTATTAAGTATGATGATGATGATGATGATGATG 551

QY 626 TGGAGAGCCCTTTTCAGGCGGCGCTGAGTGGAGTGGAGCGGCGGCGGCTGAGAGTGC 685  
DB 552 TGGAGAGCCCTTTTCAGGCGGCGCTGAGTGGAGTGGAGCGGCGGCGGCTGAGAGTGC 611

QY 686 CCATCATGTCCTCATGTCACCTCATGAGCTCTGCGCTCTACCTGCTGCTGCTGCTGCT 745  
DB 612 CCATCATGTCCTCATGTCACCTCATGAGCTCTGCGCTCTACCTGCTGCTGCTGCTGCT 671

QY 746 GGGAGTTTGGTTTGGAGGCGCATCGGCGCTTTTGTGAAGACGACGCGGCTGATCCCTCAA 805  
DB 672 GGGAGTTTGGTTTGGAGGCGCATCGGCGCTTTTGTGAAGACGACGCGGCTGATCCCTCAA 731

QY 806 GAGTCAAGGAGGCTGAGAGGAGTTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865  
DB 732 GAGTCAAGGAGGCTGAGAGGAGTTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791

QY 866 TCCTTCTGCTCTCTCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925  
DB 792 TCCTTCTGCTCTCTCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851

QY 926 TCTCAAAAGCGGAGGAGGCGCCCAAGAGAAAGCGCTGCTGCTA 968  
DB 852 TCTCAAAAGCGGAGGAGGCGCCCAAGAGAAAGCGCTGCTGCTA 894

## RESULT 5

AAH98320

ID AAH98320 standard; cDNA; 978 BP.

AC AAH98320;

XX

XX

DT 12-OCT-2001 (first entry)

XX

Human EST-derived coding sequence SEQ ID NO: 177.  
Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
diagnostics; forensic test; gene mapping; genetic disorder;  
biodiversity; gene therapy; nutrition; ss.  
Homo sapiens.  
W0200154477-A2.  
XX 02-AUG-2001.  
XX 25-JAN-2001; 2001WO-US02687.  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Wehman T;  
XX WPI; 2001-476164/51.  
XX P-PSDB; AAM23661.  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
antibodies and research use -  
XX Claim 1; Page 299-300; 1275pp; English.  
XX The present invention provides the protein and coding sequences of novel  
proteins from a variety of organisms, including human, dog, cat, horse,  
cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
urchin and tomato. These were derived from expressed sequence tags (ESTs)  
from the organism of interest. They can be used in diagnostics,  
forensics, gene mapping, identification of mutations, to assess  
biodiversity and for nutritional purposes. The present sequence is a cDNA  
of the invention.  
XX Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 other;

Query Match 68.1%; Score 858.6; DB 22; Length 978;  
Best Local Similarity 96.2%; Pred. No. 9.7e-221;  
Matches 901; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 86 GAGGCGACGTCTCGTGGTCTGAGCGCGCGGCGAGAGCGGCGGCGGCTGATCGGCT 145  
DB 5 GTGGCGACGTCTCGTGGTCTGAGCGCGCGGCGAGAGCGGCGGCGGCTGATCGGCT 64

QY 146 CCCTCGAAGTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGAGCCCGCCCGGCTCCA 205  
DB 65 CCCTCGAAGTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGAGCCCGCCCGGCTCCA 124

QY 206 AAAGCTCCAGAGGCTCCCGAGGACCGGTGCTCGGCCCTTCTTCGGTCAAGAGTGGC 265  
DB 125 AAAGCTCCAGAGGCTCCCGAGGACCGGTGCTCGGCCCTTCTTCGGTCAAGAGTGGC 184

QY 266 CCCTCGGAGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 325  
DB 185 CCCTCGGAGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 244

QY 326 TGACCGAGGAGTCTCTCTGTGTAGCTTGGAGCGCGCGAGCCCGAGAGATGCTGCTGCT 385  
DB 245 TGACCGAGGAGTCTCTCTGTGTAGCTTGGAGCGCGCGAGCCCGAGAGATGCTGCTGCT 304

QY 386 TCAATAGATGTTTCCCTCGGCTTCTCTGCTGCTTATCTACTGCTGCTGCTGCTGCT 445  
DB 305 TCAATAGATGTTTCCCTCGGCTTCTCTGCTGCTTATCTACTGCTGCTGCTGCTGCT 364

QY 446 CTGCTGTGTGAGAGTGGCCCTCGAGAGCGGAGCGGCGGCTGAGAGGCGGCGGCTGAGTGC 505





CC involved in the generation of pain and other sensory or perceptive nerve  
 CC impulses, in the establishment and endurance of mood, neurodegenerative  
 CC and sleep disorders, and in the control of muscle contraction, including  
 CC movements such as the heartbeat, digestion and vascular tone. the  
 CC sequences can be used in predictive medicine, screening and diagnostic  
 CC assays, and in pharmacogenomics.  
 XX  
 SQ Sequence 3108 BP; 830 A; 748 C; 727 G; 803 T; 0 other;  
 Query Match 44.4%; Score 559.6; DB 22; Length 3108;  
 Best Local Similarity 84.4%; Pred. No. 4.2e-140; Indels 7; Gaps 2;  
 Matches 655; Conservative 0; Mismatches 114;  
 QY 371 AGAAGATGCTCCCTCAATAGATGTTTCCCTGGCTTCTCTGCTTATCTACTGCG 430  
 Db 73 AAAAGATGCTCCCTCAACAGATGTTCCCTAGCTTCTCTAGTGCTCATCTACTGCG 132  
 QY 431 TCAGTGTCTGCTCCCTGCTGTGTGGAGTGCCTCCGAGACGAGCGCTGCAGGCA 490  
 Db 133 TCAGAGTGTCTCCCTGCTGTGTGGAGTGCCTCCGAGACGAGCGCTGCAGGCA 192  
 QY 491 ACCCATGAAGCTGCGCTGCATCTCTCTCATGAAGAGAGAGGTGAGGCCACACCG 550  
 Db 193 ATCCCATGAAGCTGAGTGCATCTCTCTCATGAAGAGAGAGGTGAGGCCACACCTG 252  
 QY 551 TGTGTGATGCTTCTACAGGCGCGAGGCGGTGAAGATTCCTTATTACAGTATCGGA 610  
 Db 253 TGTGTGATGCTTCTACAGGCGCGAGGCGGTGAAGATTCCTTATTACAGTATCGGA 312  
 QY 611 ATGGCCACAGAGGTGGAGAGCCCTTTACGGGCGCTGACGTGAAGTGCAGCAAGG 670  
 Db 313 ATGGCCACAGAGGTGGAGAGCCCTTTACGGGCGCTGACGTGAAGTGCAGCAAGG 372  
 QY 671 ACTGACAGAGTGTCCATCACTGTCTCAAGTCACTCACTCAAGTCTGCTGCTTACAG 730  
 Db 373 ACTGACAGAGTGTCCATCACTGTCTCAAGTCACTCACTCAAGTCTGCTGCTTACAG 432  
 QY 731 CTGTGATGTCTCCGGAGTTGAGTTGAGGCGCATCGCCCTTTGTGAAGACGACG 790  
 Db 433 CATGCAATGTCTCCGGAGTTGAGTTGAGGCGCATCGCCCTTTGTGAAGACGACG 492  
 QY 791 GGTGATCCCTTAAGTACACGAGGAGGCTGAGAGGACTTCACTCTGTGTCTCAG 850  
 Db 493 GACTGATACCTTTGGAGTCACTGAAGAGGCGGAGAGACTTCACTCTGTGTCTCAG 552  
 QY 851 AAATCATGATGTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 910  
 Db 553 AAATCATGATGTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 612  
 QY 911 ATGCTACAGAAAGTCTCMAAGCGGAGAGGCGCCAGAGAAACGCTCTGACTACC 970  
 Db 613 ATGCTACAGAAAGTCTCMAAGCGGAGAGGCGCCAGAGAAACGCTCTGACTACC 672  
 QY 971 TTGCCATCCATCTTGAGAACAGAGAACTCTGCGGTACCACTGAGGAATAGAACAGGA 1030  
 Db 673 TTGCTATCCCTTCAGAGACAGAGAACTCTGTGTACTCTGTGGAGAAATA-----AT 726  
 QY 1031 GAGTGTGATGAGTGGCTTGAAACATCTGAGGAGCTGAGACATCCCATGTTACGCAATG 1090  
 Db 727 GTGGTGTGACTTGGAGTATCTG-AGTGTGAGGAGCTGGATATCCCCAGTTCAGTGATG 785  
 QY 1091 TCAATGGCATCAGGAGGCGGCCCAAGGCGCCATCGCTCTCCCTTCATGATCATCAT 1146  
 Db 786 CCAGCAATATCAGGAGTGCCTCCAGGTGTCCACACATCATCTTTTCTATTCT 841

RESULT 9

AAC67836

ID AAC67836 standard; cDNA; 2220 BP.

XX

AC AAC67836;

XX

DT 15-FEB-2001 (first entry)

XX Rat beta3 subunit cDNA.  
 DE  
 XX Rat; beta sub-unit; beta3; analgesic; anticonvulsant; cerebroprotective;  
 XX vasotropic; cardiant; nootropic; cytostatic; dermatological;  
 KW gene therapy; voltage-gated sodium channel; pain; epilepsy; stroke;  
 KW ischaemia; heart disease; Jacobsen Syndrome;  
 KW familial nonchromaffin paraganglioma; phenylketonuria;  
 KW Charcot Marie Tooth disease; ss.  
 XX  
 OS Rattus sp.  
 XX  
 PN W0200063367-A1.  
 XX  
 PD 26-OCT-2000.  
 XX  
 PF 24-FEB-2000; 2000WO-EP01783.  
 XX  
 PR 15-APR-1999; 99US-0129473.  
 XX  
 PA (WARN) WARNER LAMBERT CO.  
 PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.  
 XX  
 PI Cox P, Dixon A, Jackson A, Morgan K;  
 XX  
 DR WPI; 2000-865241/64.  
 DR P-PSDB; AAB36001.  
 XX  
 PT Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
 PT sodium channel, and their corresponding polypeptides, useful for  
 PT detecting and treating sodium channel-associated conditions, e.g. pain,  
 PT epilepsy and stroke -  
 XX  
 PS Claim 6; Page 69-70; 86pp; English.  
 CC  
 CC The present sequence is given in the claims of a specification  
 CC relating to a novel family of beta sub-unit proteins from a  
 CC voltage-gated sodium channel. Human and rat beta sub-units, which  
 CC have been collectively identified as beta3, have been isolated for  
 CC the polynucleotides and polypeptides are useful for screening for  
 CC agonists and antagonists of sodium channels. The agonists, antagonists,  
 CC proteins and nucleic acids may be used diagnosing of treating diseases  
 CC or conditions associated with voltage-gated sodium channels, e.g. pain,  
 CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial  
 CC Nonchromaffin Paraganglioma, Phenylketonuria and Charcot Marie Tooth  
 CC disease.  
 XX  
 SQ Sequence 2220 BP; 573 A; 557 C; 561 G; 529 T; 0 other;  
 Query Match 42.8%; Score 540; DB 21; Length 2220;  
 Best Local Similarity 89.3%; Pred. No. 6.9e-135;  
 Matches 582; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
 QY 371 AGAAGATGCTCCCTTCAATAGATGTTTCCCTGGCTTCTCTGCTTATCTACTGCG 430  
 Db 358 AAAAGATGCTCCCTTCAACAGATGTTTCCCTAGCTTCTCTAGTCTCATCTACTGCG 417  
 QY 431 TCAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTTCGAGACGAGCGCTGCAGGCA 490  
 Db 418 TCAGAGTGTCTTCCCTGTGTGTGGAGTGCCTTCGAGACAGAGCGGTGCAGGCA 477  
 QY 491 ACCCATGAAGCTGCGCTGCATCTCTCATGAAGAGAGAGGTGAGGCCACACCG 550  
 Db 478 ATCCCATGAAGCTGAGGTGCATCTCTCTCATGAAGAGAGAGGTGAGGCCACACCTG 537  
 QY 551 TGTGTGAATGTTCTACAGGCGCCGAGGCGGTGAAGATTTCTTATTACAGTATCGGA 610  
 Db 538 TGTGTGAGTGTCTTACAGGCGCTGAGGCGGTGAAGATTTCTTATTACAGTATCGGA 597  
 QY 611 ATGGCCACAGAGGTGGAGAGCCCTTTACGGGCGCTGACGTGAAGTGCAGCAAGG 670  
 Db 598 ATGGCCACAGAGGTGGAGAGCCCTTTCCAAAGGCGCTCTGACGTGAATGGAGCAAG 657

Query Match 42.4%; Score 534.6; DB 22; Length 645;  
Best Local Similarity 89.3%; Pred. No. 1.2e-133;

RESULT 11	
AAV86895	Expressed sequence tag; secreted protein; haematopoiesis regulator;
ID AAV86895 standard; cDNA; 471 BP.	tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
XX	chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX AC	receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
AAV86895;	
27-APR-1999 (first entry)	
XX DE	EST clone BM4.
XX	
DT KW	Expressed sequence tag; secreted protein; haematopoiesis regulator;
XX KW	tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
XX KW	chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX KW	receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX	
OS Homo sapiens.	
XX	
PN W09845435-A2.	
XX	
PD 15-OCT-1998.	
XX	
PF 10-APR-1998; 98WO-US06954.	
XX	
PR 10-APR-1997; 97US-0835913.	
XX	
PA (GEM) GENETICS INST INC.	

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
 PI Racie LA, Spaulding V, Treacy W;  
 XX WPI; 1999-070076/06.  
 XX New polynucleotides encoding human secreted proteins - derived from  
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
 PT ovary, pituitary, retina and colon cDNA libraries  
 XX  
 PS Claim 1; Page 383-384; 633pp; English.  
 XX  
 CC This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene  
 CC therapy.  
 XX  
 SQ Sequence 471 BP; 93 A; 126 C; 143 G; 108 T; 1 other;

Query Match 34.0%; Score 429; DB 20; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 2.7e-105;  
 Matches 440; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 369 CCAGAGATGCTCCCTCAATAGATGTTTCCCTGGCTTCTCTGCTTACTACTG 428  
 DB 22 CCAGAGATGCTCCCTCAATAGATGTTTCCCTGGCTTCTCTGCTTACTACTG 81  
 QY 429 GGTGAGTGTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488  
 DB 82 GGTGAGTGTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140  
 QY 489 CAACCCATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 548  
 DB 141 CAACCCATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 200  
 QY 549 GGTGAGTGTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 608  
 DB 201 GGTGAGTGTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 260  
 QY 609 GAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCTGCTGCTGCTGCTG 668  
 DB 261 GAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCTGCTGCTGCTGCTG 320  
 QY 669 GGCCTGAGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728  
 DB 321 GGCCTGAGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380  
 QY 729 CACCTGCAATGTCTCCGGAGTGTGAGTTGAGTTGAGGGCCGCTGCTGCTGCTGCTG 788  
 DB 381 CACCTGCAATGTCTCCGGAGTGTGAGTTGAGTTGAGGGCCGCTGCTGCTGCTGCTG 440  
 QY 789 GCGGTGATCCCTTAAAGTTC 810  
 DB 441 GCGGTGATCCCTTAAAGTTC 462

## RESULT 12

AAS86763  
 ID AAS86763 standard; cDNA; 3531 BP.  
 XX  
 AC AAS86763;  
 XX  
 DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22567.  
 DE  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-US08631.  
 PF  
 XX 31-MAR-2000; 2000US-0540217.  
 PR  
 XX 23-AUG-2000; 2000US-0649167.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Drmanac RT, Liu C, Tang YT;  
 PI  
 XX WPI; 2001-639362/73.  
 DR  
 XX P-PDB; ABG22576.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 PT  
 XX Claim 1; SEQ ID No 22567; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 3531 BP; 952 A; 901 C; 960 G; 718 T; 0 other;  
 Query Match 31.1%; Score 392.6; DB 23; Length 3531;  
 Best Local Similarity 99.0%; Pred. No. 3.7e-95;  
 Matches 395; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 422 TCTACTGGTTCAGTGTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481  
 DB 2951 TCAGGTGTGTCAGTGTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3010  
 QY 482 TGCAGGGCAACCCCATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541  
 DB 3011 TGCAGGGCAACCCCATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3070  
 QY 542 CCACACAGCGTGTGGAATGTTCTACAGGCCCGAGGGCGGTAAGATTCTTATTACG 601  
 DB 3071 CCACACAGCGTGTGGAATGTTCTACAGGCCCGAGGGCGGTAAGATTCTTATTACG 3130  
 QY 602 AGTATCGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGGCTGCTGCTGGAATG 661

Db 3131 AGTATCGGATGCGCCACCGAGGCTGGAGAGCCCTTTTCAGGGGCGCTCGAGTGGGAAG 3190  
QY 662 GCAGCAAGGACCTGCGAGGACGTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTG 721  
Db 3191 GCAGCAAGGACCTGCGAGGACGTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTG 3250  
QY 722 GCCTCTACACCTGCAATGTGTCGCGGAGTGTGAGTTGAGGCGCATCGGCCCTTTGGA 781  
Db 3251 GCCTCTACACCTGCAATGTGTCGCGGAGTGTGAGTTGAGGCGCATCGGCCCTTTGGA 3310  
QY 782 AGACGACGCGGCTGATCCCTTAAGAGTCAACGAGGAGG 820  
Db 3311 AGACGACGCGGCTGATCCCTTAAGAGTCAACGAGGAGG 3349

## RESULT 13

AA886762  
ID AA886762 standard; cDNA; 621 BP.

XX AC AA886762;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #22566.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG22575.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

XX Claim 1; SEQ ID No 22566; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA864197-AA894564 represent novel human  
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from Wipo

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 621 BP; 156 A; 155 C; 175 G; 135 T; 0 other;

Query Match 17.8%; Score 224.4; DB 23; Length 621;  
Best Local Similarity 99.6%; Pred. No. 3.3e-50;  
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 595 ATTACGAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCGAG 654

Db 58 ATTACGAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCGAG 117

QY 655 TGGAAATGGCAGCAAGGACCTTGACGAGACGTGTCATCACTGTCTCAACGTCACTCTGAAC 714

Db 118 TGGAAATGGCAGCAAGGACCTTGACGAGACGTGTCATCACTGTCTCAACGTCACTCTGAAC 177

QY 715 GACTCTGGGCTCTTACACCTGCAATGTGTCCTCCGGAGTTTGTAGTTTGAAGCGCATCGGCC 774

Db 178 GACTCTGGGCTCTTACACCTGCAATGTGTCCTCCGGAGTTTGTAGTTTGAAGCGCATCGGCC 237

QY 775 TTTGTGAAGACGACGCGGCTGATCCCTTAAGAGTCAACGAGGAGG 820

Db 238 TTTGTGAAGACGACGCGGCTGATCCCTTAAGAGTCAACGAGGAGG 283

## RESULT 14

AAAL04971  
ID AAAL04971 standard; DNA; 4625 BP.

XX AC AAAL04971;

XX 21-NOV-2001 (first entry)

DE Human reproductive system related antigen DNA SEQ ID NO: 7659.

XX Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01339.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216847.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220563.

XX 26-JUL-2000; 2000US-0220564.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225216.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 16-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0225681.  
PR 22-AUG-2000; 2000US-0225686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0251989.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen  
XX is used in preventing, treating or ameliorating a medical condition -  
XX  
XX Disclosure; SEQ ID NO 7659; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
XX number of human reproductive system related antigens. These can be used  
XX in the prevention and treatment of reproductive system disorders,  
XX including cancer. The present sequence is a genomic sequence encoding a  
XX protein of the invention.  
XX  
XX Sequence 4625 BP; 1222 A; 1130 C; 906 G; 1367 T; 0 other;  
SQ  
Query Match 17.1%; Score 216; DB 22; Length 4625;  
Best Local Similarity 100.0%; Pred. No. 1.3e-47;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1044 AGTGGCCTGAACACCTGAGGGACCTGGACATCCCATGTTGACAAATGTCATGGCATCAG 1103  
Db 33 AGTGGCCTGAACACCTGAGGGACCTGGACATCCCATGTTGACAAATGTCATGGCATCAG 92  
Qy 1104 GAGGGCGCCCAAGGGCCCCCATCGCTTCCCTTCATGCATCCATGTTCTGTTTCATTCATT 1163  
Db 93 GAGGGCGCCCAAGGGCCCCCATCGCTTCCCTTCATGCATCCATGTTCTGTTTCATTCATT 152  
Qy 1164 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTTAACATCCATCAGACT 1223  
Db 153 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTTAACATCCATCAGACT 212  
Qy 1224 CTACGCACCATTAAGACTCTGCGCAGACTGAGAGGCC 1259

Db 213 CTAGCACCATAAGACTCTGCCAGAACTGAGAAGCC 248  
|||||

RESULT 15

ABL97864  
ID ABL97864 standard; DNA; 4625 BP.

XX ABL97864;

XX 21-JUN-2002 (first entry)

XX Human testicular antigen encoding DNA fragment SEQ ID NO: 2516.

XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
XX reproductive system disorder; urinary system disorder; gene therapy;  
XX cardiovascular disorder; respiratory disorder; neurological disorder;  
XX gastrointestinal disease; infection; cytostatic; gene; ds.

XX Homo sapiens.

XX WO200155317-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01329.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 11-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 18-AUG-2000; 2000US-0225759.

XX 22-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226688.

XX 22-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227189.

XX 30-AUG-2000; 2000US-0228924.

XX 01-SEP-2000; 2000US-0229287.

XX 01-SEP-2000; 2000US-0229343.

XX 01-SEP-2000; 2000US-0229344.

XX 01-SEP-2000; 2000US-0229345.

XX 05-SEP-2000; 2000US-0229509.

XX 05-SEP-2000; 2000US-0229513.

XX 06-SEP-2000; 2000US-0230437.

XX 06-SEP-2000; 2000US-0230438.

XX 08-SEP-2000; 2000US-0231242.

XX 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-024617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251888.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating testicular cancer -  
XX  
XX Disclosure; SEQ ID NO 2516; 766pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer.  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention.  
XX  
XX Sequence 4625 BP; 1222 A; 1130 C; 906 G; 1367 T; 0 other;  
SQ

Query Match 17.1%; Score 216; DB 23; Length 4625;  
Best Local Similarity 100.0%; Pred. No. 1.3e-47;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1044 AGGTGGCTGAACACCTGAGGACTGGACATCCCATGTTTCAGCAATGTCATGGCATCAG 1103  
Db 33 AGGTGGCTGAACACCTGAGGACTGGACATCCCATGTTTCAGCAATGTCATGGCATCAG 92

Qy 1104 GAGGGCGCCCAAGGGCCCATCGTTCCCTTCATGATCCCATGTTTCGTTTCATT 1163  
Db 93 GAGGGCGCCCAAGGGCCCATCGTTCCCTTCATGATCCCATGTTTCGTTTCATT 152

Qy 1164 CATCCATAGATCCACCTGCTCTGAGCTTTCAGCTTTCCTGACTCCCTAACTCCATCAGACT 1223  
Db 153 CATCCATAGATCCACCTGCTCTGAGCTTTCAGCTTTCCTGACTCCCTAACTCCATCAGACT 212

Qy 1224 CTAGCACCATTAGACTCTCCAGACTGAGAGCC 1259  
Db 213 CTAGCACCATTAGACTCTCCAGACTGAGAGCC 248

Search completed: December 5, 2003, 17:55:02  
Job time : 479.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:08:07 ; Search time 3728 Seconds

(without alignments)  
8221.015 Million cell updates/sec

Title: US-09-936-680-4

Perfect score: 1261  
Sequence: 1 cccctccctccgagctgagc.....tgcagaaactgagaagccgg 1261

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estm:\*

4: em\_estm:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771.8	61.2	1201	13	EX420015
2	652.6	51.8	1201	13	EX445002
3	550.6	43.7	1359	11	AK076466 Mus muscu
4	542.6	43.0	3549	11	AK049747 Mus muscu

5	531.6	42.2	4149	11	AK049286
6	524.4	41.6	672	10	BB614118
7	521.4	41.3	636	12	BM933157 UI-M-BH3-
8	513	40.7	1201	9	AL534136
9	473.6	37.6	950	10	BG294174
10	453.2	35.9	723	14	CA749311
11	447.2	35.5	927	13	BQ713131
12	397.4	31.0	1069	12	BM928131
13	390.8	31.0	825	13	BQ745919
14	386.4	30.6	846	14	CD355879
15	363.4	28.8	742	14	CB526211
16	354	28.1	845	14	CA327438
17	352.8	28.0	986	13	BU118914
18	349.2	27.7	700	12	B1739617
19	348	27.6	977	13	EX452112
20	341.8	27.1	652	10	BB652801
21	332.4	26.4	892	13	BU368614
22	326.4	25.9	595	13	BU368531
23	310	24.6	450	14	CB787935
24	307.2	24.4	712	14	CB526257
25	305.2	24.2	780	13	BQ770528
26	297.2	23.6	750	14	CD349206
27	273.4	21.7	669	12	EM951151
28	269.8	21.4	471	14	CB732717
29	264.4	21.0	705	14	CD349958
30	248.8	19.7	604	14	CB581657
31	242.4	19.2	710	10	BB653495
32	227.4	18.0	669	14	CB520894
33	214.4	17.0	443	14	D44825
34	213.8	17.0	972	13	BQ066875
35	211	16.7	683	13	BU492633
36	195.2	15.5	638	10	BB657468
37	195.2	15.5	1507	11	AK051805
38	184.8	14.7	490	13	BY247892
39	177.6	14.1	620	10	BB651905
40	177.2	14.1	600	12	BI990184
41	177.2	14.1	600	12	BI990395
42	164.6	13.1	926	13	EX411288
43	154	12.2	509	13	BX280396
44	150	11.9	319	13	BY129048
45	149.8	11.9	603	12	EM490673

## ALIGNMENTS

RESULT 1  
EX420015

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX420015 1201 bp mRNA linear EST 13-MAY-2003  
EX420015 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone  
CS0DF03YA09 5-PRIME, mRNA sequence.

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015

EX420015



435 GAGGAGGTCGAGCCACCGCGTGTGGAATGGTTCTACAGCCCGGCGGTAAAGAT 494  
589 TTCTCTTATTTACGATATCGGAATGCGCCACGAGAGTGGAGAGCCCTTTTCAGGCGCGC 648  
495 TTCTCTTATTTACGATATCGGAATGCGCCACGAGAGTGGAGAGCCCTTTTCAGGCGCGC 554  
649 CTGCGATGGAATGCGAGAGAGCTGCGAGAGCTGCTCATCTGCTGCTCAAGTCACT 708  
555 CTGCGATGGAATGCGAGAGAGCTGCGAGAGCTGCTCATCTGCTGCTCAAGTCACT 614  
709 CTGAACAGCTCTGGCTCTTACACCTGCTGATGTGCTCCCGGAGTGTGATTTGAGCGGCAT 768  
615 CTGAACAGCTCTGGCTCTTACACCTGCTGATGTGCTCCCGGAGTGTGATTTGAGCGGCAT 674  
769 CGGCGCTTTGGAAGAGAGCGCGGTGATCCCTTACAGAGTACCGAGAGGCTTGGAGAG 828  
675 CGGCGCTTTGGAAGAGAGCGCGGTGATCCCTTACAGAGTACCGAGAGGCTTGGAGAG 734  
829 GACTTCACCTCTGTGCTCTCAAGATCATGATGATCATCTCTGCTTCTCTCCTCAGCCTG 888  
735 GACTTCACCTCTGTGCTCTCAAGATCATGATGATCATCTCTGCTTCTCTCCTCAGCCTG 794  
889 TGCTGCTCATGAGATGATATTTGCTTACAGAGAGTCTCAAGAGCGAGAGCGCGCC 948  
795 TGCTGCTCATGAGATGATATTTGCTTACAGAGAGTCTCAAGAGCGAGAGCGCGCC 854  
949 CAAGAAAGCGCTCTGACTTACCTTCCATCTGAGAAACAGAGAGTCTCTCGGTA 1008  
855 CAAGAAAGCGCTCTGACTTACCTTCCATCTGAGAAACAGAGAGTCTCTCGGTA 913  
1009 CCAGTGAGAGATAGACAGGAG 1031  
914 MCAGTGGAGCATARAGAGGAG 936  
RESULT 3  
AK076466  
LOCUS  
DEFINITION  
AK076466 1359 bp mRNA linear HTC 07-DEC-2002  
Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched  
library, clone:4833414B02 product:VOLTAGE-GATED SODIUM CHANNEL  
BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog (Rattus  
norvegicus), full insert sequence.  
ACCESSION  
AK076466.1 GI:26345409  
VERSION  
AK076466.1  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1  
Carninci, P. and Hayashizaki, Y.  
TITLE  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL  
MEDLINE  
99279253  
PUBMED  
10349636  
REFERENCE  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
JOURNAL  
MEDLINE  
20499374  
PUBMED  
11042159  
REFERENCE  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunaga, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORSGenome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Balzarini, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Fonzald, M.F., Brownstein, M.J., Bult, C., Hill, D.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Rongwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S.  
and Hayashizaki, Y.  
TITLE  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
JOURNAL  
MEDLINE  
21085660  
PUBMED  
11217851  
5TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORSThe PANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1359)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Aizawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,  
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,  
Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,  
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,  
Nunuma, K., Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,  
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, I.,  
Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,  
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORSSubmitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORSSubmitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .1359  
/organism="Mus musculus"  
/mol\_type="cDNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:4833414B02"  
/db\_xref="MGI:1904202"  
/db\_xref="taxon:10090"  
/clone="4833414B02"  
/tissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"FEATURES  
sourcecDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .1359  
/organism="Mus musculus"  
/mol\_type="cDNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:4833414B02"  
/db\_xref="MGI:1904202"  
/db\_xref="taxon:10090"  
/clone="4833414B02"  
/tissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"

CDS  
 /dev stage="0 day neonate"  
 57...704  
 /notes="unnamed protein product; VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus] (SPTR|Q9J300, evidence: PASTY, 99.5%ID, 100%length, match=646)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAC36356.1"  
 /db\_xref="GI:26345410"  
 /translation="MPAFNRLPLASLVLYVVRVCFVCPVEPSETRAVQNSMKLR CISCRRVEATVYVWFYRPEFGKDFLYEYRNGHOEVSFPQRLQWMSKDLQD VSTLVNLNDSGLYTCNVSFEFPEARFPVKTRLIPLRVTEAGEDTSTVSEI MMVILAVFLTLMFLFIEMIVCYRKVSKABEAQENASDYLAIPESENKENVVPVEE"  
 BASE COUNT 343 a 342 c 314 g 360 t  
 ORIGIN  
 Query Match 43.7%; Score 550.6; DB 11; Length 1359;  
 Best Local Similarity 87.2%; Pred. No. 3.7e-128;  
 Matches 604; Conservative 0; Mismatches 89; Indels 0; Gaps 0;  
 Qy 332 AGGGAATCTCTGCTAGCTTGGAGCGCCAGCCAGAGATGCTGCTTCAATA 391  
 Db |||||  
 Qy 392 GATTGTTCCCTGGCTTCTCTGCTTATCTACTGCGGTCAGTGTCTCTTCCCTGTGT 451  
 Db |||||  
 Qy 73 GATTGTTCCCTAGCTTCTAGTGTCTACTTACTGCTGCTGCTGCTGCTGCTGCT 132  
 Db |||||  
 Qy 452 GTGTGAGTCCCTCGAGACGGGCGCTGCGAGGCGACCCAGCTGAGCTGCGTCA 511  
 Db |||||  
 Qy 133 GTGTGAGTACCTTCGGACAGAGCGCTGCGAGGCGAATTCATGAAGCTGAGATCA 192  
 Db |||||  
 Qy 512 TCTCCTCATGAAGAGAGAGAGTGGAGCGCCAGCGTGTGGAAATGGTCTACAGGC 571  
 Db |||||  
 Qy 193 TCTCCTCATGAAGAGAGAGTGGAGCGCCAGCGTGTGGAAATGGTCTACAGGC 252  
 Db |||||  
 Qy 572 CGAGGCGGTAAAGATTCTTATTTACAGTATCGAATGCGCCAGCGAGGTGAGCA 631  
 Db |||||  
 Qy 253 CTGAGGCGGTAAAGATTCTTATTTATGAGTATCGAATGCGCCAGCGAGGTGAGCA 312  
 Db |||||  
 Qy 632 GCCCTTTCAGGGCGGCTCGAGTGAATGGCAGCAGGACCTGCGAGCGTGTCCATCA 691  
 Db |||||  
 Qy 313 GCCCTTTCAGGGCGGCTCGAGTGAATGGCAGCAGGACCTGCGAGCGTGTCCATCA 372  
 Db |||||  
 Qy 692 CTGTGCTCAGCTCAGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751  
 Db |||||  
 Qy 373 CTGTGCTCAGCTCAGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432  
 Db |||||  
 Qy 752 TTGAGTTTGAAGGCGCATCGGCGCTTTGTGAAGACGAGCGGCTGATCCCTTAAGATCA 811  
 Db |||||  
 Qy 433 TTGAGTTTGAAGGCGCATCGGCGCTTTGTGAAGACGAGCGTATACCTCCGAGTCA 492  
 Db |||||  
 Qy 812 CGAGGAGGCTGAGAGGAGCTTCACTCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 871  
 Db |||||  
 Qy 493 CTGAGGAGGCGGAGAGAGCTTCACTCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 552  
 Db |||||  
 Qy 872 TGGTCTTCTCACCTGCTGCTGCTCATCGAGATGATATTTCTACAGAAAGGCTCTCAA 931  
 Db |||||  
 Qy 553 TGGTCTTCTCACCTGCTGCTGCTTTATTTAGATGATATTTCTACAGAAAGGCTCTCAA 612  
 Db |||||  
 Qy 932 AAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 991  
 Db |||||  
 Qy 613 AGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 672  
 Db |||||  
 Qy 992 AGGAGAGCTCTGGGTACCGAGTGGAGGAGTGA 1024  
 Db |||||  
 Qy 673 AGGAGAGCTCTGGGTACCGAGTGGAGGAGTGA 705  
 Db |||||

RESULT 4  
 AK049747

LOCUS  
 DEFINITION  
 AK049747 3549 bp mRNA linear HTC 05-DEC-2002  
 Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length enriched library, clone:CS30046E12 product:VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus], full insert sequence.  
 ACCSSION  
 VERSION AK049747.1 GI:26340475  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 99279253  
 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 11042159  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861  
 4  
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hata, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stabili, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaudo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Righwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schönbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohtsuki, S. and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851  
 5  
 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 3549)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, H., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohashi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers

## FEATURES

1..3549

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:CS30046E12"  
/db\_xref="taxon:10090"  
/clone="CS30046E12"  
/tissue types="spinal cord"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_tag="12 days embryo"  
232..879  
/note="unnamed protein product; VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus] (SPTR|Q9JUK0, evidence: PASTY, 99.5%ID, 100%length, match=646)  
putative"

## CDS

/codon\_start=1  
/protein\_id="BAC33901.1"  
/db\_xref="GI:26340476"  
/translation="MFANRLPLSLVLYVWVFCVVPVETPAVQNSKLR  
CISCKEEVEATVVEFYRPGKDFLYEYRNGHVEVFPQGLRWGSKDLQ  
VSTLVNLTSLGLYCNVRFEPFEPFVFKTRLPVLTPEAGEDFTSVVSEI  
MWTLVFLTLWLFIBMIYCYRKYRKAEBAAQENASDYLAIPISENKENVVPEE"  
BASE COUNT 924 a 850 c 825 g 950 t

## ORIGIN

Query Match 43.0%; Score 542.6; DB 11; Length 3549;  
Best Local Similarity 88.2%; Pred. No. 6,1e-126;  
Matches 590; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 356 GAAGCGGCGACCCAGAGATGCTGCTTCAATAGATGTTTCCCTGGCTTCTCTCG 415  
DB 212 GAACCACTTACGAGAAAGATGCTGCTTCAACAGATGCTTCCCTAGCTTCTCTAG 271

QY 416 TGCTTATCTACTGGGTGAGTGTGCTGCTTCCCTGTGTGTGGAGTGCCTTCGGAGCGG 475  
DB 272 TGCTCATCTACTGGGTGAGTGTGCTGCTTCCCTGTGTGTGGAGTGCCTTCGGAGCGG 331

QY 476 AGGCGGTGAGGCAACCCATGAAGTGTGCTGCTTCCCTCATGAAGAGAGAGAGG 535  
DB 332 AAGCGGTGAGGCAACCCATGAAGTGTGCTGCTTCCCTCATGAAGAGAGAGAGG 391

QY 536 TGGAGGCGACACCGGTGAGTGTGCTGCTTCCCTCATGAAGAGAGAGAGGTTCTT 595  
DB 392 TGGAGGCGACACCGGTGAGTGTGCTGCTTCCCTCATGAAGAGAGAGAGGTTCTT 451

QY 596 TTACGAGTATCGGAATGGCCACCGAGGAGTGGAGAGCGCCCTTTCAGGGGCGCTGCACT 655

Db 452 TATATGAGTATCGAATGGCCACCGAGGAGTGGAGAGCCCTTCCAGGTGCTGCACT 511  
QY 656 GGAATGGCAGCAGGACCTCGAGGACGCTGCTCATCTGCTCAACGTCACCTCTGAACG 715  
DB 512 GGRATGGGAGCAAGACCTCGAGGAGCTATCCATCCTGTTCTCAATGTCACCTCTGAATG 571  
QY 716 ACTCTGGCTCTACACCTGCTCATGTCCTCCGGAGTGTGAGTTGAGCGGCGATCGGCT 775  
DB 572 ACTCTGGCTCTACACCTGCTCATGTCCTCCGGAGTGTGAGTTGAGCGGCGATCGGCT 631  
QY 776 TTGTGAGCAGCAGCGGCTGATCCCTTAAGAGTCCACGAGGAGCTCGAGAGGACTTCA 835  
DB 632 TTGTGAGCAGCAGGACTATACCTCCGAGTCTAGAGGCGGAGAGACTTCA 691  
QY 836 CTTCTGTGCTCAGAAATCATGATGATCATCTCTGCTGCTTCTCTCATCCCTGCTGCTG 895  
DB 692 CTTCTGTGCTCAGAAATCATGATGATCATCTCTGCTGCTTCTCTCATCCCTGCTGCTG 751  
QY 896 TCATCCAGATGATATATGCTACAGAAAGTCTCAAAAGCCGAGAGGAGCGCCCAAGAAA 955  
DB 752 TTATTTGAGTATCTATGCTACAGAAAGTCTCTAAGCCGAGAGGCGGAGCTCAGGAAA 811  
QY 956 AGCGGCTCAGTACCTTGCCTATCCATCTGAGAGACAGAGAGAACTCTCGGTACAGTGG 1015  
DB 812 ATCGGCTCAGTACCTTGCCTATCCATCTGAGAGACAGAGAGAACTCTCTGTACCCGCTG 871

1016 AGGAATAGA 1024  
872 AGGAATAGA 880

RESULT 5  
AK049286  
LOCUS  
DEFINITION  
Mus musculus ES cells cDNA, RIKEN full-length enriched library,  
clone: C330019103 product: VOLTAGE-GATED SODIUM CHANNEL BETA-3  
SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus  
norvegicus], full insert sequence.  
AK049286  
VERSION AK049286.1 GI:26093400  
KEYWORDS HTG; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
MEDLINE 10349636  
PUBMED 10349636

REFERENCE  
2  
Carninci, P., Shibata, Y., Hayate, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20493174  
MEDLINE 11042159  
PUBMED 11042159

REFERENCE  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
MEDLINE 11076861  
PUBMED 11076861

REFERENCE  
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamazaki, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schiraldi, L. M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wnshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851

REFERENCE  
AUTHORS

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 4149)

JOURNAL  
NEDLINE  
PUBMED  
REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE  
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES  
source

Location/Qualifiers  
 1. .4149  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="PANTOM DB:C330019103"  
 /db\_xref="taxon:10090"  
 /clone="C330019103"  
 /cell\_type="ES cells"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 410. .1058  
 /note="VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog (Rattus norvegicus) (SPR1Q9JK00, evidence: PASTY, 99.5%id, 100%length,

## misc\_feature

match=646)  
 putative"  
 4129. .4134  
 /note="putative"  
 4149  
 /notes="putative"  
 BASE COUNT 1088 a 1022 c 961 g 1077 t 1 others  
 ORIGIN

Query Match 42.2%; Score 531.6; DB 11; Length 4149;  
 Best Local Similarity 88.1%; Pred. No 4e-123;  
 Matches 590; Conservative 0; Mismatches 79; Indels 1; Gaps 1;

QY 356 GARAGCCGCCAGCCCCAGAGATCCCTGCGCTCAATAGATGTTTCCCTCGCTTCG 415  
 DB 390 GAAGACCACTTACGAGAAAGATCCCTGCGCTCAACAGATGCTTCCCTAGCTTAG 449  
 QY 416 TGCTTACTACTGGTGCAGTGTCTGCTTC-CCTGCTGTGTGAGAGTCCCTCGGAGAG 474  
 DB 450 TGCTTACTACTGGTGCAGTGTCTGCTTCNCTGTGTGTAGAGTACCTCGGAGACA 509  
 QY 475 GAGGCCGTGACGGGCAACCCATGAAGCTGCGCTGCATCTCCTGCAATGAAGAGAGAG 534  
 DB 510 GAAGCCGTGACGGGCAATTCATGAAGCTGAGATGCATCTCCTGCAATGAAGAGAGAG 569  
 QY 535 GTGCGGCCACCAACGCTGCTGGAATGTTCTACAGCCCGAGGCGGTAAAGATTTCTT 594  
 DB 570 GTGAGGCCACCACTGCTAGTGGAGTGGTCTACAGGCCCTGAGGCGGTAAAGATTTCTT 629  
 QY 595 ATTTACAGTATCGGAATGCGCACACAGGAGGTGAGAGAGGCCCTTTTCAGGGCGGCTTCGAG 654  
 DB 630 ATATATGAGTATCGAATGCGCACACAGGAGGTGAGAGAGGCCCTTTTCAGGGCGGCTTCGAG 689  
 QY 655 TGGATGGCAGCAGGACCTGCGAGGACCTGCGATCACTGCTGCTGCACTCACTCACTCACT 714  
 DB 690 TGGATGGCAGCAGGACCTGCGAGGACCTGCGATCACTGCTGCTGCACTCACTCACTCACT 749  
 QY 715 GACTCTGCTCTACACCTTCAATGTTGTCGGGAGTGTGAGTTTCAGGCGCATCGGCC 774  
 DB 750 GACTCTGCTCTACACCTTCAATGTTGTCGGGAGTGTGAGTTTCAGGCGCATCGGCC 809  
 QY 775 TTGTGAGAGCAGCGCGCTGATCCCTTAAGAGTCAAGAGAGGCTCGAGAGAGGCTCGAGAGAGCTTC 834  
 DB 810 TTGTGAGAGCAGCGCGCTGATCCCTTAAGAGTCAAGAGAGGCTCGAGAGAGGCTCGAGAGAGCTTC 869  
 QY 835 ACCTCTGCTCTCAGAAATCATGATGATCATCTTCTGCTCTTCTCACTCACTCGCTGCGTG 894  
 DB 870 ACCTCTGCTCTCGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 929  
 QY 895 CTGATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
 DB 930 TTATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989  
 QY 955 AACCGCTCTCACTACCTTGCATCCATCCATCTGAGAAACAGGAGAACTCTCGGATACAGTG 1014  
 DB 990 AATCGCTCTCACTACCTTGCATCCATCCATCTGAGAAACAGGAGAACTCTCGGATACAGTG 1049

QY 1015 GAGGAATAGA 1024  
 DB 1050 GAGGAATAAA 1059

## RESULT 6

BB614118  
 LOCUS BB614118 RIKEN full-length enriched, 0 day neonate head Mus  
 DEFINITION musculus cDNA clone 4833414B02 5', mRNA sequence.  
 ACCSSION BB614118  
 VERSION BB614118  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: MEST@mail.nih.gov  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 REVERSE.

# FEATURES

source  
1. 636  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-bseq-G-04-0-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH BMAP M S4"  
/notes="Vector: pTV73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The  
NIH BMAP M S4 library is a subtracted library of a series,  
ultimately derived from a mixture of individually tagged,  
normalized libraries from ten regions of the mouse brain  
(cerebellum, brain stems, olfactory bulbs, hypothalamus,  
cortex, amygdala, basal ganglia, pineal gland, striatum,  
hippocampus) after a series of subtractions to reduce the  
representation of cDNAs from which ESTs had already been  
generated. The following serially subtracted libraries  
were generated in this process: NIH BMAP M S4,  
NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,  
NIH BMAP M S2, NIH BMAP M S1. The subtracted library  
(NIH BMAP M S4) was constructed as follows: PCR amplified  
cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and  
NIH BMAP M S3.1 clones from which 3' ESTs had been derived  
was used as a driver in a hybridization with a pool of  
the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1  
libraries in the form of single-stranded circles. The  
remaining single-stranded circles (subtracted library)  
was purified by hydroxyapatite column chromatography,  
converted to double-stranded circles and electroporated  
into DH10B bacteria (Life Technologies) to generate the  
NIH BMAP M S4 library. This procedure has been previously  
described (Bonaldo, Lennon and Soares, Genome Research  
6:791-806, 1996)."

BASE COUNT 158 a 154 c 169 g 155 t  
ORIGIN  
Query Match 41.3%; Score 521.4; DB 12; Length 636;  
Best Local Similarity 88.8%; Pred. No. 6.3e-121;  
Matches 564; Conservative 0; Mismatches 71; Indels 0; Gaps 0;  
QY 371 AGAAGATGCTGCTTCATAGATGTTTCCCTGGCTCTCTCGTGCTATCTACTGGG 430  
DB 2 AAAAGATGCTGCTTCACAGATTGCTTCCCTAGCTTCTTAGTGCTCATCTACTGG 61  
QY 431 TCAGTGTCTGCTTCCCTGTGTGTGTGAGATGTCCTCGAGACGAGGCCGCTGCAGGCA 490  
DB 62 TCAGAGTCTGCTTCCCTGTGTGTGTAGATGATACCTTCGAGACAGAACCGCTGCAGGCA 121  
QY 491 ACCCATGAAGTGGCTGCTATCTCTCGATGAGAGAGAGAGGTGGAGCCACACGG 550  
DB 122 ATTCATGAGCTGAGATGATCTCTGTCATGAGAGAGAGAGGTGGAGCCACCATG 181  
QY 551 TGTGTGAATGGTCTTACAGGCCCGCGGTAAGATTCCTTTATTTACGATATCGGA 610

DB 182 TAGTGGAGTGGTCTACAGGCTGAGGGCGTAAAGATTTCCTATATAGTATCGAA 241  
QY 611 ATGGCCACAGGAGGTGGAGGCCCTTTCAGGGGGGCTGCAGTGGATGCGACGAAGG 670  
DB 242 ATGGCCACAGGAGGTGGAGGCCCTTTCAGGGGGGCTGCAGTGGATGCGACGAAG 301  
QY 671 ACCTGAGGACGTGTCTCATCTGTGCTCAACGTCCTCTGAAACGACTCTGSCCTCTACA 730  
DB 302 ACCTGAGGACGTGTCTCATCTGTGCTCAACGTCCTCTGAAACGACTCTGSCCTCTACA 361  
QY 731 CCTGCAATGTGTCGGGAGTTTGTGATTTGAGGCGCATCGGCCCTTGTGAGACGACG 790  
DB 362 CATGTATGTGTCAGGGAGTTTGTGATTTGAGGCGCATCGGCCCTTGTGAGACGAC 421  
QY 791 GGCTGATCCCCCTTAAGAGTCACGAGGAGGTGGAGAGGACTTCACTCTGTGGTCTCAG 850  
DB 422 GACTATATCCCTGCGAGTCACTGAAGAGCGGGAGAGACTTCACTCTCGGTCTCGG 481  
QY 851 AATCATGATGATCATCTCTGTGGTCTTCTCCTCACCTGTGGTCTCATFCGAGATGATAT 910  
DB 482 AATCATGATGATCATCTCTGTGGTCTTCTCCTCACCTGTGGTCTCATFCGAGATGAT 541  
QY 911 ATTGCTACAGAAAGGTCTCAAAAGCGAAGAGCGCAGCCCAAGAAAACGGCTCTGACTACC 970  
DB 542 ATTGCTACAGAAAGGTCTCAAAAGCGAAGAGCGCAGCCCAAGAAAACGGCTCTGACTACC 601  
QY 971 TTGCCATCCATCTGAGAACACAGGAGAACTCTGGG 1005  
DB 602 TTGCTATCCCTTCAGAGAACACAGGAACTCTGTG 636

## RESULT 8

## LOCUS

## DEFINITION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

AL534136 1201 bp mRNA linear EST 12-MAY-2003  
AL534136 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
CS0DF005YI02 5-PRIME, mRNA sequence.

AL534136  
EST.

AL534136.2 GI:30539643

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

On Feb 13, 2001 this sequence version replaced gi:12797629.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 6147.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DF005BE01Q1&cluster=6147.r. Contact :

Feng Liang Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0DF005BE01Q1.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DF005YI02"

/tissue\_type="FETAL BRAIN"

/dev\_stage="fetal"

/clone\_lib="Homo sapiens FETAL BRAIN"

/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

Email: [cgapbs-remail.nin.gov](mailto:cgapbs-remail.nin.gov)  
Tissue Procurement: The Cepko Laboratory

5 ATGCGTCTGACTACTT 680

## RESULT 10

CA749311 723 bp mRNA linear EST 27-NOV-2002  
 LOCUS UI-M-FYO-cdd-k-17-0-UI.r1 NIH\_BMAP\_FYO Mus musculus cDNA clone  
 DEFINITION IMAGE: 6831594 5', mRNA sequence.  
 CA749311  
 ACCSSION CA749311.1 GI:25570984  
 VERSION  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: csapbs-remail.nih.gov  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Cloning Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 723

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE: 6831594"

/tissue\_type="whole brain"

/dev\_stage="embryo 13.5, 14.5, 15.5, 16.5, 17.5dpc"

/lab\_host="DH10B (TI phage resistant)"

/clone\_lib="NIH\_BMAP\_FYO"

/notes="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is AGCAGACAG. This library was created for the University

of Iowa Brain Anatomy Project (BMAP); 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,

program coordinator."

Program coordinator."

176 a 179 c 201 g 165 t 2 others

## BASE COUNT

## ORIGIN

Query Match 35.9%; Score 453.2; DB 14; Length 723;  
 Best Local Similarity 88.0%; P-adj. No. 1.2e-103;  
 Matches 505; Conservative 0; Mismatches 68; Indels 1; Gaps 1;  
 QY 356 GAAGCGCCAGCCCGAAGATGCTTCCCTTCAATAGATGTTTCCCTCGGCTTCTCTCG 415  
 DB 150 GAACACCACTTACGAGAAAGATGCTTCCCTTCAACAGATGCTTCCCTAGCTTCTCTAG 209  
 QY 416 TGCTTATCTACTGGTCTAGTGTCTGCTTCCCTGTTGTGTGGAGTCCCTCGGACGCG 475  
 DB 210 TGCTTATCTACTGGTCTAGTGTCTGCTTCCCTGTTGTGTGGAGTCCCTCGGACGCG 269  
 QY 476 AGGCGGTGACGGGCAACCCCATGAAGTGGCTGCTATCTCTCTCATGAAGAGAGAGGAGG 535  
 DB 270 AAGCCGTGCAGGGCAATTCATGAAGCTGAGATGCTCTCTCATGAAGAGAGAGGAGG 329

QY 536 TGGAGGCCACACCGTGGTGAATGTTCTACAGGCCCGAGGCGGTAAAGATTTCTTTA 595  
 DB 330 TGGAGGCCACACCTAGTGGAGTGTCTACAGGCTGAGGCGGTAAAGATTTCTTTA 389  
 QY 596 TTTACGAGTATCGGAATGGCCACAGAGGTGGAGGCCCTTTTACGGGCGCTGCTGAGT 655  
 DB 390 TATATGAGTATCGAAATGGCCACAGAGGTGGAGGCCCTTTTCCAAAGGTGCTGCTGAGT 449  
 QY 656 GBAATGCCAGCAGGACCTGCGAGAGGTGTCATCACTGTCACAGTCACTCTCAACG 715  
 DB 450 GBAATGGGAGCAAGACCTGCGAGAGGTGTCATCACTGTCACAGTCACTCTCAACG 509  
 QY 716 ACTTGGCCCTTACACCTGCAATGTGTCCGGGAGTTGAGTTGAGGCGCATTCGCCCT 775  
 DB 510 ACTTGGCCCTTACACATGTAATGTGTCCAGGAGTTGAGTTGAGGCGCATTCGCCCT 569  
 QY 776 TTGTGAAGACGACGCGCTGATCCCTTAAAGTGTACCG-AGAGGCTGGAGGACTTC 834  
 DB 570 TTGTGAAGACGACGACTAATACCTTGGAGTCACTGAAGAAGCGGGAGAGACTTC 629  
 QY 835 ACCTTCTGTGTCAGAAATCATGATGATCATCTTCTGCTTCTTCTCACCTTGGGCTG 894  
 DB 630 ACCTTCTGTGTCAGAAATCATGATGATCATCTTCTGCTTCTTCTCACCTTGGGCTG 689  
 QY 895 CTATTCAGATGATATATTCCTACAGAAAGTCT 928  
 DB 690 TTTATTGAGATGATCTATTCTTACAGAAAGTCT 723

## RESULT 11

## BQ713131

## LOCUS

## DEFINITION

## 5', mRNA sequence.

## ACCESSION

## BQ713131

## VERSION

## BQ713131.1

## KEYWORDS

## EST.

## SOURCE

## Mus musculus (house mouse)

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## Contact: Robert Strausberg, Ph.D.

## Email: csapbs-remail.nih.gov

## Tissue Procurement: Susan L. Sullivan, PhD.

## cDNA Library Preparation: ResGen, Invitrogen Corp

## cDNA Library Arrayed by: The I.M.A.G.E. Consortium

## DNA Sequencing by: Agencourt Bioscience Corporation

## Clone distribution: MGC clone distribution information can be

## found through the I.M.A.G.E. Consortium/LLNL at:

## http://image.llnl.gov

## Plate: LLAM13727 row: b column: 23

## High quality sequence stop: 679.

## Location/Qualifiers

## 1. 927

## /organism="Mus musculus"

## /mol\_type="mRNA"

## /db\_xref="taxon:10090"

## /clone="IMAGE:6308278"

## /lab\_host="DH10B (phage-resistant)"

## /clone\_lib="NIH\_MGC 129"

## /notes="SPOR6:1.cdd; Site:1: EcoRV; Site:2: NotI; Cloned

## unidirectionally. Primer: Oligo dT. Average insert size

## 2.2 Kb. Constructed by ResGen, Invitrogen Corp. Note: this

## is a NIH\_MGC Library."

## BASE COUNT

## ORIGIN

## 218 a 241 C 258 g 210 t

```

http://image.llnl.gov
Plate: LLAMI2832 row: f column: 04
High quality sequence stop: 592.
Location/Qualifiers
1. .1069
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5769987"
  /lab_host="DH10B"
  /clone_lib="NIH MGC 121"
  /note="Organ: brain; Vector: pCMV-SPORT6; Site.1: NotCl;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned [EcoRV site is destroyed upon
cloning]. Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH MGC Library."
200 a 376 c 302 g 191 t

```

Query Match	31.5%; Score 397.4; DB 12; Length 1069;
Best Local Similarity	89.9%; Pred. No. 1.8e-89;
Matches	461; Conservative 0; Mismatches 46; Indels 6; Gaps 3;
QY	349 AGCCTGGAGCGCCAGCCAGAGATGCTCCCTCAATAGATGTTCCCCCTGCGT 408
DB	
QY	355 AGCCTGGAGCGCCAGCCAGAGATGCTCCCTCAATAGATGTTCCCCCTGCGT 414
DB	
QY	409 TCTCTGCTGTTATCTATCTGGGTGAGTGTCTGCTCCCTGTGTGTGGAGTGCCCTCG 468
DB	
QY	415 TCTCTGCTGTTATCTATCTGGGTGAGTGTCTGCTCCCTGTGTGTGGAGTGCCCTCG 474
DB	
QY	469 GAGACGGAGCGCGTGCAAGGCAACCCCATGAAGTGGCTGCATCTCTCGATGAAGAGA 528
DB	
QY	475 GAGACGGAGCGCGTGCAAGGCAACCCCATGAAGTGGCTGCATCTCTCGATGAAGAGA 534
DB	
QY	529 GAGGAGTGGAGGCCACACGCTGGTGGATGGTCTACAGGCCGAGGCGGTAAAGAT 588
DB	
QY	535 GAGGAGTGGAGGCCACACGCTGGTGGATGGTCTACAGGCCGAGGCGGTAAAGAT 594
DB	
QY	589 TTCCTTTATTTAGAGTATCGGAATGGCCACACGAGGTGGAGAGCCCTTTTCAGGGGCGC 648
DB	
QY	595 TTCCTTTATTTAGAGTATCGGAATGGCCACACGAGGTGGAGAGCCCTTTTCAGGGGCGC 654
DB	
QY	649 CTGCACTGGAATGGCAGCAAGGACCTGCAGAGCGTGTCCATCACTGTGTCTCAACGTCACT 708
DB	
QY	655 CTGCACTGGAATGGCAGCAAGGACCTGCAGAGCGTGTCCATCACTGTGTCTCAACGTCACT 714
DB	
QY	709 CTGAACGACTCTGGCCCTTACACCTGCATGTGTCCCGGGAGTTTTCAGTATT---GAGGGCG 765
DB	
QY	715 CTGAACGATCTGGCCCTTACACCTGCATGTGTCCCGGGAGTTTTCAGTATT---GAGGGCG 774
DB	
QY	766 CATCGGCCCTTTGTGAAGA-CGACGGCGGTGATTCGCCCTTAAGAGTCAACCGAGGAGGC--T 822
DB	
QY	775 CCCCCTCCCTTCGTGGAACCCACCGGGCTGATTCCTCTACATACACCCAGGAGGCCCTG 834
DB	
QY	823 GGAAGGACTTCACTCTGTGTCTCAGAAATC 855
DB	
QY	835 GAAAGGCTTCACTCTGAAAGGCCTCCGAAAC 867
DB	

RESULT 13	BQ745919	825 bp	mRNA	linear	EST 17-JUL-2002
LOCUS	UI-N-EXO-b2k-d-18-0-UI.r1 NIH	EMAP_EXO	Mus musculus	cDNA clone	
DEFINITION	IMAGE:5709281	5', mRNA sequence.			
ACCESSION	BQ745919				
VERSION	BQ745919.1	GI:21892693			
KEYWORDS	EST.				

RESULT 12	1069 bp	mRNA	linear	EST 12-MAR-2002		
BM928131	AGNCOURT_6699645	NIH_MGC_121	Homo sapiens	cDNA clone IMAGE:5769887		
LOCUS	5', mRNA sequence.					
DEFINITION						
ACCESSION	BM928131					
VERSION	BM928131.1	GI:19378510				
KEYWORDS	EST.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens (human)					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE	1 (bases 1 to 1069)					
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> ,					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)					
	Unpublished					
	Contact: Robert Strausberg, Ph.D.					
	Email: <a href="mailto:cgapbs-re@mail.nih.gov">cgapbs-re@mail.nih.gov</a>					
	Tissue procurement: Life Technologies, Inc.					
	cDNA Library Preparation: Life Technologies, Inc.					
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
	DNA sequencing by: Agencourt Bioscience Corporation					
	Clone distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E Consortium/LLNL at:					

**SOURCE**  
Mus musculus (house mouse)  
**ORGANISM**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE**  
1 (bases 1 to 825)  
NIH-MGC <http://mgc.nci.nih.gov/>  
**AUTHORS**  
National Institutes of Health, Mammalian Gene Collection (MGC)  
**TITLE**  
Unpublished  
**COMMENT**  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

**FEATURES**  
source  
Seq primer: PYX-5.  
Location/Qualifiers  
1..825  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5709281"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 15.5 dpc"  
/lab\_host="DH10B (TI phage resistant)"  
/clone\_lib="NIH-BMAP-Exo"  
/note="Organ: brain; vector: PYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured mRNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with an  
oligo-AT primer containing a Not I site. Double stranded  
cDNA was size selected according to mRNA size fraction.  
ligated with EcoR I adaptor, digested with Not I, and then  
cloned directionally into PYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is GTCCTGGGAA. This library was created for the  
University of Iowa Mouse Brain Molecular Anatomy Project  
(BMAP): 'Gene Discovery in the Developing Mouse Nervous  
System', supported by National Institutes of Mental Health  
(NIMH), Hemin Chin, Ph.D., program coordinator."  
186 a 226 c 232 g 178 t 3 others

**BASE COUNT**  
ORIGIN  
Query Match 31.0%; Score 390.8; DB 13; Length 825;  
Best Local Similarity 87.2%; Pred. No. 7.4e-88;  
Matches 472: Conservative 0; Mismatches 65; Indels 4; Gaps 4;  
QY 356 GAAGCGCCGACCCGAGAGATGCTGCTTCAATAGATGTTTCCCTGGCTTCTCTCG 415  
DB 278 GAACACCTTACGAGAAAGATGCTGCTTCAACAGATGCTTCCCTAGTCTCTAG 337  
QY 416 TGCTTATCTATGGGTGAGTGTGCTTCCCTGTGTGTGAGAGTGCCTTCGGAGCGG 475  
DB 338 TGCTCATCTACTGGGTGAGATGCTGCTTCCCTGTGTGTGAGAGTGCCTTCGGAGCAG 397  
QY 476 AGGCGCTGACGGGCAACCCATGAAGTGGCTCATCTCTCATGAGAGAGGAGCG 535  
DB 398 AGCCGTGACGGGCATTCATGAGCTGAGATGCTTCTCTGATGAGAGGAGGAGG 457  
QY 536 TCGAGGCCACACCGGTGTGGAATGGTTTACAGCCCGAGGGCGGTAAGATTCTCTTA 595  
DB 458 TCGAGGCCACACCTGATGAGGTGTTTCTACAGCCCTGAGGGCGGTAAAGATTCTCTTA 517  
QY 596 TTTCAGGTATCGGAATGSCCACCAGAGGTGAGAGCCCTTTCAGGGGCGCTCGAGT 655  
DB 518 TATATGAGTATCGAAATGGCCACCAGGAGGTGGAGAGCCCTTTCAGAGGTCTGTCAGT 577

QY 656 GGAATGGCAGCAGGACCTGCGAGGAGTGTCCATCAGTGTCTCAAGCTCACTCTGAACG 715  
DB 578 GGAATGGGAGGAGAAAGACCTGCGAGGAGTATCATCAGCTGTTCTCAATGTCACTCTGAATG 637  
QY 716 ACTCTGGCCTCTACACCTGCAATGTGTCGCGGAGTTTGAAGTGAAGCGGATCGGCCT 775  
DB 638 ACTCTGGCCTCTACACATGTAATGTGTCAGNGAGTTTGAAGTGAAGCACACCGGCCCT 697  
QY 776 TTGTGAGAGCAGCGGCTGATCCCTTAAGAGTACACGAGGAGGCTGGAGAGGACTTCA 835  
DB 698 TTGTGAGAGC-ACAGAGACTATATCCCTGCGAGTCTCTGAAGAGCGGAGAGACTTCA 756  
QY 836 CTTCTGTGTCRCAGAAATCATGATGATCATCTCTCTGTTCTTCTCACTCTGTCGTC 895  
DB 757 -CTCCGTGCTTC-GNAATCATGATGATCATCTCTCTGTC-TCCTCACTTGTGGCTGT 813  
QY 896 T 896  
DB 814 T 814

**RESULT 14**  
CD355879  
**LOCUS**  
DEFINITION  
UI-M-FY0-cgo-1-21-0-UI.r1 NIH-BMAP\_FY0 Mus musculus cDNA clone  
IMAGE: 30355580 5', mRNA sequence.  
CD355879  
CD355879.1 GI:31148380  
EST.  
**KEYWORDS**  
Mus musculus (house mouse)  
**ORGANISM**  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE**  
1 (bases 1 to 846)  
NIH-MGC <http://mgc.nci.nih.gov/>  
**AUTHORS**  
National Institutes of Health, Mammalian Gene Collection (MGC)  
**TITLE**  
Unpublished  
**JOURNAL**  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
**COMMENT**  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mouseefl.html>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

**FEATURES**  
source  
Seq primer: PYX-5.  
Location/Qualifiers  
1..846  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30355580"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (TI phage resistant)"  
/clone\_lib="NIH-BMAP\_FY0"  
/note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into PYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is AGAGGAGCAG. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:21:16 ; Search time 111 Seconds  
(without alignments)  
5014.271 Million cell updates/sec

Title: US-09-936-680-4  
Perfect score: 1261  
Sequence: 1 cccctccctccagctgagc.....tgcagaaactgagaagccgg 1261

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44.8	3.6	11558	5	PCT-US93-06251-23
2	38.6	3.1	1941	3	US-09-082-737-1
3	37.2	3.0	152331	3	US-09-128-155-16
4	36.4	2.9	1658	3	US-09-385-028-20
5	36.4	2.9	1658	4	US-09-726-614-20
6	36.4	2.9	11604	3	US-09-385-028-13
7	36.4	2.9	11604	4	US-09-726-614-13
8	36.4	2.9	15079	3	US-09-385-028-1
9	36.4	2.9	15079	4	US-09-726-614-1
10	36.2	2.9	1371	4	US-09-996-243-388
11	36.2	2.9	2608	4	US-09-904-615-16
12	36.2	2.9	7218	1	US-08-232-463-14
13	36.2	2.9	38506	3	US-09-320-878-19
14	36.2	2.9	38506	4	US-09-141-908-1
15	36.2	2.9	38506	4	US-09-657-440-19
16	36	2.9	1020	4	US-09-252-991A-198
17	35.8	2.8	1236	4	US-09-423-439-9
18	35.8	2.8	1317	4	US-09-252-991A-3363
19	35.8	2.8	1342	3	US-08-445-515-52
20	35.8	2.8	1342	3	US-08-445-515-53
21	35.8	2.8	1758	4	US-09-252-991A-3311
22	35.8	2.8	1926	4	US-09-423-439-25
23	35.8	2.8	1929	4	US-09-423-439-15
24	35.8	2.8	1995	4	US-09-252-991A-3322
25	35.8	2.8	1998	4	US-09-423-439-50
26	35.8	2.8	2019	4	US-09-423-439-31
27	35.8	2.8	2025	4	US-09-423-439-37

Sequence 52, Appl  
Sequence 104, Appl  
Sequence 2593, Ap  
Sequence 2325, Ap  
Sequence 1, Appl  
Sequence 22, Appl  
Sequence 22, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 2512, Ap  
Sequence 9, Appl  
Sequence 9, Appl  
Sequence 23, Appl  
Sequence 23, Appl

28 35.8 2.8 3217 4 US-09-423-439-52  
29 35.8 2.8 3521 4 US-09-484-970B-104  
30 35.6 2.8 1851 4 US-09-252-991A-2593  
31 35.6 2.8 2172 4 US-09-252-991A-2325  
32 35.6 2.8 6152 1 US-08-557-139-1  
33 35.4 2.8 1994 3 US-08-600-982-22  
34 35.4 2.8 1994 5 PCT-US94-10261A-22  
35 35.4 2.8 2769 3 US-09-118-408-1  
36 35.4 2.8 2769 4 US-09-506-855-1  
37 35.4 2.8 2769 4 US-09-911-1768-1  
38 35.4 2.8 2769 4 US-09-619-740-1  
39 35.4 2.8 2769 4 US-09-506-852-1  
40 35.4 2.8 47981 4 US-09-679-279-1  
41 35.2 2.8 894 4 US-09-252-991A-2512  
42 35.2 2.8 2681 1 US-08-070-165F-9  
43 35.2 2.8 2681 2 US-08-885-418-9  
44 35.2 2.8 5496 3 US-08-600-982-23  
45 35.2 2.8 5496 5 PCT-US94-10261A-23

## ALIGNMENTS

## RESULT 1

PCT-US93-06251-23

Sequence 23, Application PC/TUS9306251

GENERAL INFORMATION:

APPLICANT: Wickstrom, Eric and Rife, Jason P.

TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06251

FILING DATE: 19930630

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Digiclio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8586

TELECOMMUNICATION INFORMATION:

TELEPHONE: 516-742-4343

TELEFAX: 516-742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 11558 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US93-06251-23

Query Match 3.6%; Score 44.8; DB 5; Length 11558;

Best Local Similarity 56.7%; Pred. No. 0.045; Indels 1; Gaps 1;

Matches 102; Conservative 0; Mismatches 77; Indels 1; Gaps 1;

QY 74 GGGGTGGCGGGGAGGCGACTGTCTCGTGTCTGAGCGCGGAGAGCGGCGCGAGC 133

Db 428 GGGAGGGCGGGTAGGGACCTTTTCAGAGCCAGAGGGCTTTTCGGGGCGCGCGC 487

QY 134 GCGTGTATCGGTCCCTCGACTCGAGGAGTCCAGTGGGGTCTG-CTTAGGGCCCAAGCCC 192



APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDRA UR  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-385-028-20

Query Match 2.9%; Score 36.4; DB 3; Length 1668;  
Best Local Similarity 47.4%; Pred. No. 3.7;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 618 CCAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCACTGGAATGGCAGAGGACCTGCA 677  
DB 1050 CTTGACGGCTACAGCACTTCGACCGCTACCGGCTCGGCCCGAGGGGACCGGGACCT 1109  
QY 678 GGAGGTGTCATCTGCTCAAGTCACTCTGACGACTCTGCTCTACCTCTACCTGCAA 737  
DB 1110 GGAGGCGCGCCGCGCGAGCTGAAGCTGCGCGGGATGCGCGGCTTCGCGCACCAAGAT 1169  
QY 738 TGTGTCCCGGAGTTTGAATTTGAGCGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGAT 797  
DB 1170 CGCGCGCCGCAAGGACCGGCTCAAGAGTACCGGCGCGCGGCGCTGGCGCGGGCT 1229  
QY 798 CCCCTTAAGATCAGGAGGCTCGAGAGGCTTCACTCTGTGTCT 847  
DB 1230 CGCGCGGCTCGCATCGAGGCGGAGGTCTGAGCTTCCCGTGGCGGACT 1279

RESULT 5  
US-09-726-614-20  
Sequence 20, Application US/09726614  
Patent No. 6514735  
GENERAL INFORMATION:  
APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A. Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6514735  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jennifer Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/726,614  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price

REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDRA UR  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-726-614-20

Query Match 2.9%; Score 36.4; DB 4; Length 1668;  
Best Local Similarity 47.4%; Pred. No. 3.7;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 618 CCAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCACTGGAATGGCAGAGGACCTGCA 677  
DB 1050 CTTGACGGCTACAGCACTTCGACCGCTACCGGCTCGGCCCGAGGGGACCGGGACCT 1109  
QY 678 GGAGGTGTCATCTGCTCAAGTCACTCTGACGACTCTGCTCTACCTCTACCTGCAA 737  
DB 1110 GGAGGCGCGCCGCGCGAGCTGAAGCTGCGCGGGATGCGCGGCTTCGCGCACCAAGAT 1169  
QY 738 TGTGTCCCGGAGTTTGAATTTGAGCGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGAT 797  
DB 1170 CGCGCGCCGCAAGGACCGGCTCAAGAGTACCGGCGCGCGGCGCTGGCGCGGGCT 1229  
QY 798 CCCCTTAAGATCAGGAGGCTCGAGAGGCTTCACTCTGTGTCT 847  
DB 1230 CGCGCGGCTCGCATCGAGGCGGAGGTCTGAGCTTCCCGTGGCGGACT 1279

RESULT 6  
US-09-385-028-13  
Sequence 13, Application US/09385028  
Patent No. 6232106  
GENERAL INFORMATION:  
APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A. Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6232106  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jennifer Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/385,028  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELEPHONE: (202) 638-6666

TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDEA UR  
INFORMATION FOR SEQ ID NO: 13;  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11604 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-385-028-13

Query Match 2.9%; Score 36.4; DB 3; Length 11604;  
Best Local Similarity 47.4%; Pred. No. 7.4;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 618 CCAGGAGGTGGAGAGCCCTTTTCAGGGGGCCCTGTCAGTGGATGCGAGCAGCACTGCA 677  
DB 8258 CCTCGAGCGGTACAAGCACTTCGACCGCTACCGGTGCGGCCGAGGGCACCGGCACT 8317  
QY 678 GGAGGTGTCATCACTGTGCTCAACGTCACTCTGAAGCACTCTGGCTCTACACTGCAA 737  
DB 8318 GGAGGGCGCCCGCGGAGCTGAAGCTGCGCGGATGCCCCGACGGCTTCCGACCAAGAT 8377  
QY 738 TGTGTCTCCCGGAGTTGAGTTTCAGGCGCATCGGCCCTTTGTGAAGACGACGCGCTGAT 797  
DB 8378 CGCCGCGCGAGGACCGGCTCAAGAGTAGTACCGGCGCGGAGGCGCTGGCCCGCGGCT 8437  
QY 798 CCCCTTAAGAGTACCGAGAGGCTGGAGAGGACTTCACTCTGTGTGCTCT 847  
DB 8438 CGCCGCGGTGCGGATCGAGCGGAGGTGCTGGACTTCCGTCGGGCGACT 8487

## RESULT 7

US-09-726-614-13  
Sequence 13, Application US/09726614  
Patent No. 6514735  
GENERAL INFORMATION:  
APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6514735  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jennifer Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/726,614  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 39305350  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDEA UR  
INFORMATION FOR SEQ ID NO: 13;  
SEQUENCE CHARACTERISTICS:

LENGTH: 11604 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-726-614-13

Query Match 2.9%; Score 36.4; DB 4; Length 11604;  
Best Local Similarity 47.4%; Pred. No. 7.4;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 618 CCAGGAGGTGGAGAGCCCTTTTCAGGGGGCCCTGTCAGTGGATGCGAGCAGCACTGCA 677  
DB 8258 CCTCGAGCGGTACAAGCACTTCGACCGCTACCGGTGCGGCCGAGGGCACCGGCACT 8317  
QY 678 GGAGGTGTCATCACTGTGCTCAACGTCACTCTGAAGCACTCTGGCTCTACACTGCAA 737  
DB 8318 GGAGGGCGCCCGCGGAGCTGAAGCTGCGCGGATGCCCCGACGGCTTCCGACCAAGAT 8377  
QY 738 TGTGTCTCCCGGAGTTGAGTTTCAGGCGCATCGGCCCTTTGTGAAGACGACGCGCTGAT 797  
DB 8378 CGCCGCGCGAGGACCGGCTCAAGAGTAGTACCGGCGCGGAGGCGCTGGCCCGCGGCT 8437  
QY 798 CCCCTTAAGAGTACCGAGAGGCTGGAGAGGACTTCACTCTGTGTGCTCT 847  
DB 8438 CGCCGCGGTGCGGATCGAGCGGAGGTGCTGGACTTCCGTCGGGCGACT 8487

## RESULT 8

US-09-385-028-1  
Sequence 1, Application US/09385028  
Patent No. 6232106  
GENERAL INFORMATION:  
APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6232106  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jennifer Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/385,028  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 39305350  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDEA UR  
INFORMATION FOR SEQ ID NO: 1;  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15079 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

T-870-C0C-60-20

Query Match	2.9%;	Score 36.4;	DB 4;	Length 15079;
Best Local Similarity	47.4%;	Prod.No. 8.1;		
Matches 109;	Conservative 0;	Mismatches 121;	Indels 0;	
618	CCAGGAGCTGAGAGACCCCTTTTCAGGGCGCGCTGCAGTGTGAATGSCAGACAAGSGG			
10290	CCTCGACGGCTACAAGCACTTTCGACCGCTACCGGTGCGCCCCGAGGSCACCGAG			
678	GGACGTGTCCATCACTGTGCTCAACGTCACCTCTGNAACACTCTTGGCCTCTACACA			
10350	GGAGSCGCCCGCGCGGAGCTGTGAGCTGCGCGGAGTCCCGACCGCTTTCGGCAG			
738	TGTGTCCCGGGAGTTTGAGTTTGAGGGCGGCATCGGCCCTTTTGTGAAGACAGAGCGC			
10410	CGCCGCCCGGAGAGGACCGGCTCAAGGATGACGGGCGCCGAGGCGCTGTGGCGG			
798	CCCCCTAAGATCATCCGAGAGGCGTGTGAGAGCACTTCACTCTCTGTGTGTCT 84			
10470	CGCCCGGCTCGGCATTCGAGGGCGAGTGCTGTGACTTTCGCTGGCGGCGACT 10			

RESULT 10  
US-09-996-243-388  
; Sequence 388, Application US/09996243  
; Patent No. 6478825  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.

```

1  APPLICANT: zhang, zemin
2
3  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
4  TITLE OF INVENTION: Acids Encoding the Same
5  FILE REFERENCE: P2730P1C13
6  CURRENT APPLICATION NUMBER: US/09/9996,243
7  CURRENT FILING DATE: 2001-11-14
8  PRIOR APPLICATION NUMBER: 60/049787
9  PRIOR FILING DATE: 1997-06-16
10 PRIOR APPLICATION NUMBER: 60/062250
11 PRIOR FILING DATE: 1997-10-17
12 PRIOR APPLICATION NUMBER: 60/065186
13 PRIOR FILING DATE: 1997-11-12
14 PRIOR APPLICATION NUMBER: 60/065311
15 PRIOR FILING DATE: 1997-11-13
16 PRIOR APPLICATION NUMBER: 60/066770
17 PRIOR FILING DATE: 1997-11-24
18 PRIOR APPLICATION NUMBER: 60/075945
19 PRIOR FILING DATE: 1998-02-25
20 PRIOR APPLICATION NUMBER: 60/078910

```

1	PRIOR FILING DATE: 1998-03-20	2	PRIOR APPLICATION NUMBER: 60/089599
2	PRIOR APPLICATION NUMBER: 60/083322	3	PRIOR FILING DATE: 1998-06-17
3	PRIOR FILING DATE: 1998-04-28	4	PRIOR APPLICATION NUMBER: 60/089600
4	PRIOR APPLICATION NUMBER: 60/084600	5	PRIOR FILING DATE: 1998-06-17
5	PRIOR FILING DATE: 1998-05-07	6	PRIOR APPLICATION NUMBER: 60/089653
6	PRIOR APPLICATION NUMBER: 60/087106	7	PRIOR FILING DATE: 1998-06-17
7	PRIOR FILING DATE: 1998-05-28	8	PRIOR APPLICATION NUMBER: 60/089801
8	PRIOR APPLICATION NUMBER: 60/087607	9	PRIOR FILING DATE: 1998-06-18
9	PRIOR FILING DATE: 1998-06-02	10	PRIOR APPLICATION NUMBER: 60/089907
10	PRIOR APPLICATION NUMBER: 60/087609	11	PRIOR FILING DATE: 1998-06-18
11	PRIOR FILING DATE: 1998-06-02	12	PRIOR APPLICATION NUMBER: 60/089908
12	PRIOR APPLICATION NUMBER: 60/087759	13	PRIOR FILING DATE: 1998-06-18
13	PRIOR FILING DATE: 1998-06-02	14	PRIOR APPLICATION NUMBER: 60/089947
14	PRIOR APPLICATION NUMBER: 60/087827	15	PRIOR FILING DATE: 1998-06-19
15	PRIOR FILING DATE: 1998-06-03	16	PRIOR APPLICATION NUMBER: 60/089948
16	PRIOR APPLICATION NUMBER: 60/088021	17	PRIOR FILING DATE: 1998-06-19
17	PRIOR FILING DATE: 1998-06-04	18	PRIOR APPLICATION NUMBER: 60/089952
18	PRIOR APPLICATION NUMBER: 60/088025	19	PRIOR FILING DATE: 1998-06-19
19	PRIOR FILING DATE: 1998-06-04	20	PRIOR APPLICATION NUMBER: 60/090246
20	PRIOR APPLICATION NUMBER: 60/088026	21	PRIOR FILING DATE: 1998-06-22
21	PRIOR FILING DATE: 1998-06-04	22	PRIOR APPLICATION NUMBER: 60/090252
22	PRIOR APPLICATION NUMBER: 60/088028	23	PRIOR FILING DATE: 1998-06-22
23	PRIOR FILING DATE: 1998-06-04	24	PRIOR APPLICATION NUMBER: 60/090254
24	PRIOR APPLICATION NUMBER: 60/088029	25	PRIOR FILING DATE: 1998-06-22
25	PRIOR FILING DATE: 1998-06-04	26	PRIOR APPLICATION NUMBER: 60/090349
26	PRIOR APPLICATION NUMBER: 60/088030	27	PRIOR FILING DATE: 1998-06-23
27	PRIOR FILING DATE: 1998-06-04	28	PRIOR APPLICATION NUMBER: 60/090355
28	PRIOR APPLICATION NUMBER: 60/088033	29	PRIOR FILING DATE: 1998-06-23
29	PRIOR FILING DATE: 1998-06-04	30	PRIOR APPLICATION NUMBER: 60/090429
30	PRIOR APPLICATION NUMBER: 60/088326	31	PRIOR FILING DATE: 1998-06-24
31	PRIOR FILING DATE: 1998-06-04	32	PRIOR APPLICATION NUMBER: 60/090431
32	PRIOR APPLICATION NUMBER: 60/088167	33	PRIOR FILING DATE: 1998-06-24
33	PRIOR FILING DATE: 1998-06-05	34	PRIOR APPLICATION NUMBER: 60/090435
34	PRIOR APPLICATION NUMBER: 60/088202	35	PRIOR FILING DATE: 1998-06-24
35	PRIOR FILING DATE: 1998-06-05	36	PRIOR APPLICATION NUMBER: 60/090444
36	PRIOR APPLICATION NUMBER: 60/088212	37	PRIOR FILING DATE: 1998-06-24
37	PRIOR FILING DATE: 1998-06-05	38	PRIOR APPLICATION NUMBER: 60/090445
38	PRIOR APPLICATION NUMBER: 60/088217	39	PRIOR FILING DATE: 1998-06-24
39	PRIOR FILING DATE: 1998-06-05	40	PRIOR APPLICATION NUMBER: 60/090472
40	PRIOR APPLICATION NUMBER: 60/088655	41	PRIOR FILING DATE: 1998-06-24
41	PRIOR FILING DATE: 1998-06-09	42	PRIOR APPLICATION NUMBER: 60/090535
42	PRIOR APPLICATION NUMBER: 60/088734	43	PRIOR FILING DATE: 1998-06-24
43	PRIOR FILING DATE: 1998-06-10	44	PRIOR APPLICATION NUMBER: 60/090540
44	PRIOR APPLICATION NUMBER: 60/088738	45	PRIOR FILING DATE: 1998-06-24
45	PRIOR FILING DATE: 1998-06-10	46	PRIOR APPLICATION NUMBER: 60/090542
46	PRIOR APPLICATION NUMBER: 60/088742	47	PRIOR FILING DATE: 1998-06-24
47	PRIOR FILING DATE: 1998-06-10	48	PRIOR APPLICATION NUMBER: 60/090557
48	PRIOR APPLICATION NUMBER: 60/088810	49	PRIOR FILING DATE: 1998-06-24
49	PRIOR FILING DATE: 1998-06-10	50	PRIOR APPLICATION NUMBER: 60/090676
50	PRIOR APPLICATION NUMBER: 60/088824	51	PRIOR FILING DATE: 1998-06-25
51	PRIOR FILING DATE: 1998-06-10	52	PRIOR APPLICATION NUMBER: 60/090678
52	PRIOR APPLICATION NUMBER: 60/088826	53	PRIOR FILING DATE: 1998-06-25
53	PRIOR FILING DATE: 1998-06-10	54	PRIOR APPLICATION NUMBER: 60/090690
54	PRIOR APPLICATION NUMBER: 60/088858	55	PRIOR FILING DATE: 1998-06-25
55	PRIOR FILING DATE: 1998-06-11	56	PRIOR APPLICATION NUMBER: 60/090694
56	PRIOR APPLICATION NUMBER: 60/088861	57	PRIOR FILING DATE: 1998-06-25
57	PRIOR FILING DATE: 1998-06-11	58	PRIOR APPLICATION NUMBER: 60/090695
58	PRIOR APPLICATION NUMBER: 60/088876	59	PRIOR FILING DATE: 1998-06-25
59	PRIOR FILING DATE: 1998-06-11	60	PRIOR APPLICATION NUMBER: 60/090696
60	PRIOR APPLICATION NUMBER: 60/089105	61	PRIOR FILING DATE: 1998-06-25
61	PRIOR FILING DATE: 1998-06-12	62	PRIOR APPLICATION NUMBER: 60/090862
62	PRIOR APPLICATION NUMBER: 60/089440	63	PRIOR FILING DATE: 1998-06-26
63	PRIOR FILING DATE: 1998-06-15	64	PRIOR APPLICATION NUMBER: 60/090863
64	PRIOR APPLICATION NUMBER: 60/089512	65	PRIOR FILING DATE: 1998-06-26
65	PRIOR FILING DATE: 1998-06-16	66	PRIOR APPLICATION NUMBER: 60/091360
66	PRIOR APPLICATION NUMBER: 60/089514	67	PRIOR FILING DATE: 1998-07-01
67	PRIOR FILING DATE: 1998-06-16	68	PRIOR APPLICATION NUMBER: 60/091478
68	PRIOR APPLICATION NUMBER: 60/089532	69	PRIOR FILING DATE: 1998-07-02
69	PRIOR FILING DATE: 1998-06-17		





```

Db      3333 CCAGCCGCGTCTTCCGGAACTCCGAGATCCGCTCCCGCGCGCCACCCAGCCGCA 3274
QY      213 CCAGGCGCTTCCCGAGGACCGGTGCTCGGCGCTTCTTCGGTCAAGTCCGCGCCCTGG 272
Db      3273 CAGGTCTTCCGCGAGGACCCACCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3214
QY      273 GGGCAGTTGTCCTCCAAAGGTTTCTCGAAGAACTCTGAGAGGCGGCGGCGGCGGCGG 332
Db      3213 CGGCTGTCGACGCGCACCGGCGCGCGCCACCGGAGCGGCTCCGACCGCGCGGTTGCG 3154
QY      333 GCGAATCTCTGTGTAGCTTGGAAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 373
Db      3153 CGCGCGCTCCCGTGCACGACGAGGAGGAGCTGCTCCGCGA 3113

```

Search completed: December 5, 2003, 22:59:14  
 Job time : 116 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:39:02 ; Search time 525 Seconds  
(without alignments)  
7982.969 Million cell updates/sec

Title: US-09-936-680-4  
Perfect score: 1261  
Sequence: 1 ccctccctccgagctgagc.....tgcagaaactgagaagccgg 1261

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 166179959 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications NA:\*\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1261	100.0	1261	10	US-09-997-579-4
2	559.6	44.4	2632	13	US-10-029-191-22
3	559.6	44.4	3108	13	US-10-029-191-1
4	540	42.8	2220	10	US-09-997-579-3
5	534.6	42.4	645	13	US-10-029-191-21
6	216	17.1	4625	11	US-09-764-891-7659
7	148.2	11.8	657	13	US-10-029-191-23
8	138.6	11.0	1490	10	US-09-917-800A-1654
9	49	3.9	243	12	US-10-029-386-16214
10	49	3.9	569	12	US-10-029-386-2514
11	41.2	3.3	922	12	US-10-017-161-1435
12	41	3.3	1097	10	US-09-771-161A-58
13	41	3.3	3382	10	US-09-771-161A-59
14	40.2	3.2	2806	11	US-09-291-417-102
15	40.2	3.2	2838	12	US-10-134-102-3
16	39.8	3.2	392	10	US-09-960-352-2238

17	39.2	3.1	867	12	US-10-140-472-20	Sequence 20, Appl
18	39.2	3.1	867	12	US-10-141-761-20	Sequence 20, Appl
19	39.2	3.1	867	12	US-10-142-885-20	Sequence 20, Appl
20	39.2	3.1	867	12	US-10-158-790-20	Sequence 20, Appl
21	39.2	3.1	867	12	US-10-137-871-20	Sequence 20, Appl
22	39.2	3.1	867	12	US-10-140-805-20	Sequence 20, Appl
23	39.2	3.1	867	12	US-10-140-864-20	Sequence 20, Appl
24	39.2	3.1	867	12	US-10-140-923-20	Sequence 20, Appl
25	39.2	3.1	867	12	US-10-141-756-20	Sequence 20, Appl
26	39.2	3.1	867	12	US-10-141-759-20	Sequence 20, Appl
27	39.2	3.1	867	14	US-10-123-155-20	Sequence 20, Appl
28	39.2	3.1	867	15	US-10-146-731-20	Sequence 20, Appl
29	39.2	3.1	5973	12	US-10-094-466-63	Sequence 63, Appl
30	39.2	3.1	6202	12	US-10-120-988-120	Sequence 120, App
C 31	39	3.1	587	12	US-10-029-386-22980	Sequence 22980, A
32	39	3.1	936	10	US-09-974-300-836	Sequence 836, App
C 33	38.6	3.1	550	12	US-10-029-386-24925	Sequence 24925, A
34	38.6	3.1	1179	9	US-09-815-242-7698	Sequence 7698, Ap
C 35	38.6	3.1	2310	12	US-10-120-988-142	Sequence 142, App
36	38.4	3.0	3163	12	US-10-017-161-1857	Sequence 1857, App
37	38	3.0	1689	12	US-10-102-622-9	Sequence 9, Appli
38	38	3.0	1689	12	US-10-102-622-11	Sequence 11, Appl
C 40	38	3.0	2307	11	US-09-893-519A-87	Sequence 87, Appl
C 43	38	3.0	27079	12	US-10-034-650-55	Sequence 55, Appl
41	38	3.0	38360	11	US-09-999-121-14	Sequence 14, Appl
C 42	37.8	3.0	1091	12	US-10-027-632-30772	Sequence 30772, A
C 43	37.8	3.0	1091	13	US-10-027-632-30772	Sequence 30772, A
44	37.8	3.0	1327	12	US-10-027-632-215913	Sequence 215913,
45	37.8	3.0	1327	13	US-10-027-632-215913	Sequence 215913,

#### ALIGNMENTS

#### RESULT 1

US-09-997-579-4  
; Sequence 4, Application US/09997579  
; Patent No. US20020113203A1  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge University Technical Services  
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sc  
; TITLE OF INVENTION: channel  
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses  
; FILE REFERENCE: 674558-2001  
; CURRENT APPLICATION NUMBER: US/09/997,579  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: PCT/EP00/01783  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60,129,473  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1261  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-997-579-4

Query Match	100.0%;	Score 1261;	DB 10;	Length 1261;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1261;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	CCCTCCCTCCGAGCTGAGCTTACCCTGGCGCGCAACAGCGAGCGCGAGTGG	60	
Db	1	CCCTCCCTCCGAGCTGAGCTTACCCTGGCGCGCAACAGCGAGCGCGAGTGG	60	
QY	61	AAGCTGAGTTCGGGGCTGGCGGAGCGAGCTGCTCGCTGCTGCTGAGCGCGCGAGA	120	
Db	61	AAGCTGAGTTCGGGGCTGGCGGAGCGAGCTGCTCGTGTGCTGAGCGCGCGAGA	120	
QY	121	GGCGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTCGCTTAG	180	
Db	121	GGCGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTCGCTTAG	180	

181 GGCCTCAAGCCGCCACCCGGCTCCAAAGCTCCAGGCGCTCCCGACGACCGGCTGCTCG 240  
181 GGCCTCAAGCCGCCACCCGGCTCCAAAGCTCCAGGCGCTCCCGACGACCGGCTGCTCG 240  
241 GGCCTTCTTGGTTCAGAAAGTGGCCCTCGGGGCGAGTTCGTCCTCCAAAGGCTTCTCG 300  
241 GGCCTTCTTGGTTCAGAAAGTGGCCCTCGGGGCGAGTTCGTCCTCCAAAGGCTTCTCG 300  
301 AAAGAAATCTGAGAGGCGCAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGC 360  
301 AAAGAAATCTGAGAGGCGCAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGC 360  
361 GCGCAGCCCGCAGAGATGCTGCTTCAATAGATGTTTCCCTCGGCTTCTCTGCTGCTT 420  
361 GCGCAGCCCGCAGAGATGCTGCTTCAATAGATGTTTCCCTCGGCTTCTCTGCTGCTT 420  
421 ATCTACTGGTTCAGTGTCTGCTTCCCTGTGTGTGGAAGTCCCTCGGAGCGGAGGCC 480  
421 ATCTACTGGTTCAGTGTCTGCTTCCCTGTGTGTGGAAGTCCCTCGGAGCGGAGGCC 480  
481 GTGAGAGGCAACCCATGAAGCTGGCTGCAATCTCTGCAATGAAGAGAGAGGAGTGGAG 540  
481 GTGAGAGGCAACCCATGAAGCTGGCTGCAATCTCTGCAATGAAGAGAGAGGAGTGGAG 540  
541 GGCACACGGTGGTGGATGCTTCTACAGGCGCGAGGCGGTAAAGATTTCTTATTTAC 600  
541 GGCACACGGTGGTGGATGCTTCTACAGGCGCGAGGCGGTAAAGATTTCTTATTTAC 600  
601 GAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAATGAAT 660  
601 GAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAATGAAT 660  
661 GGCACACGGTGGTGGATGCTTCTACAGGCGCGAGGCGGTAAAGATTTCTTATTTAC 720  
661 GGCACACGGTGGTGGATGCTTCTACAGGCGCGAGGCGGTAAAGATTTCTTATTTAC 720  
721 GGCCTCTACCTGCAATGTGTCGGGAGTTTGAAGTGGAGGCGCATCGGCCCTTTGTG 780  
721 GGCCTCTACCTGCAATGTGTCGGGAGTTTGAAGTGGAGGCGCATCGGCCCTTTGTG 780  
781 AGAGCAGCGCGCTGATCCCCCTAAGAGTACCGAGGAGGCTGGAGAGGACTTCACTCT 840  
781 AGAGCAGCGCGCTGATCCCCCTAAGAGTACCGAGGAGGCTGGAGAGGACTTCACTCT 840  
841 GTGGTCTCAGAAATCATGATGTACATCTTCTGTGCTTCTCACCCTGTGGCTCATC 900  
841 GTGGTCTCAGAAATCATGATGTACATCTTCTGTGCTTCTCACCCTGTGGCTCATC 900  
901 GAGATGATATTTCTACAGAAAGTCTCAAAAGCGGAGGCGGAGGCGGAGGAGGAGGCG 960  
901 GAGATGATATTTCTACAGAAAGTCTCAAAAGCGGAGGCGGAGGCGGAGGAGGAGGCG 960  
961 TGTGACTACCTTGGCATCCCAATCTGAGAACAGAGAGACTCTGGCGGTACCAAGTGGAGAA 1020  
961 TGTGACTACCTTGGCATCCCAATCTGAGAACAGAGAGACTCTGGCGGTACCAAGTGGAGAA 1020  
1021 TAGAACAGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGAGACATCCCATG 1080  
1021 TAGAACAGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGAGACATCCCATG 1080  
1081 TTCAGCAATGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTCATGC 1140  
1081 TTCAGCAATGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTCATGC 1140  
1141 ATCCATGTTCTGTTCAATTCATTCATATCCATCCATCCATCCATCCATCCATCCATCC 1200  
1141 ATCCATGTTCTGTTCAATTCATTCATATCCATCCATCCATCCATCCATCCATCCATCC 1200  
1201 GACTCCCTAACCTCAGACCTCTACGACCATTAAGACTCTGCGAGAACTGAGAGCGG 1260  
1201 GACTCCCTAACCTCAGACCTCTACGACCATTAAGACTCTGCGAGAACTGAGAGCGG 1260

QY 1261 G 1261  
Db 1261 G 1261

## RESULT 2

US-10-029-191-22  
; Sequence 22, Application US/10029191  
; Publication No. US20020160453A1  
; GENERAL INFORMATION:  
; APPLICANT: CURTIS, ROY A. J.  
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 210147.00XX/5U1  
; CURRENT APPLICATION NUMBER: US/10/029.191  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 09/569,978  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/134,198  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 2632  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-10-029-191-22

Query Match 44.4%; Score 559.6; DB 13; Length 2632;  
Best Local Similarity 84.4%; Pred. NO. 3.8e-160;  
Matches 655; Conservative 0; Mismatches 114; Indels 7; Gaps 2;

QY 371 AGAAGATGCTGCTCCCTTCAATAGATGTTTCCCTGGCTTCTCTGCTGCTTACTTCTGGG 430  
Db 73 AAAGATGCTGCTCCCTTCAAGATGCTTCCCTAGCTTCTCTAGTCTCATCTACTGGG 132  
QY 431 TCAGTGTCTGCTTCCCTGTGTGTGGAAGTGCCTTCGAGACCGAGGCGCTGAGGCGCA 490  
Db 133 TCAGATGCTGCTTCCCTGTGTGTGGAAGTGCCTTCGAGACCGAGGCGCTGAGGCGCA 192  
QY 491 ACCCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 550  
Db 193 ATCCCATGAGCTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252  
QY 551 TGSTGGAATGTTCTACAGGCGCTGAGGCGGCTTAAAGATTTCTTATTTACGATATCGGA 610  
Db 253 TGSTGGAATGTTCTACAGGCGCTGAGGCGGCTTAAAGATTTCTTATTTACGATATCGGA 312  
QY 611 ATGGCCACCGAGGCTGAGGAGGCGCTTTCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 670  
Db 313 ATGGCCACCGAGGCTGAGGAGGCGCTTTCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 372  
QY 671 ACCTGACGAGCTGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 730  
Db 373 ACCTGACGAGCTGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432  
QY 731 CTGCAATGTTCCCGGAGTTTGAATTTGAGGCGCATCGGCGCTTCTGCTGCTGCTGCTGCTG 790  
Db 433 CATGCAATGTTCCCGGAGTTTGAATTTGAGGCGCATCGGCGCTTCTGCTGCTGCTGCTGCTG 492  
QY 791 GCTGATGCTCCCTTAAAGTCACTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTG 850  
Db 493 GACTGATGCTTTCGAGTCACTGAGGAGGCTGAGGAGGCTGAGGAGGCTGAGGAGGCTG 552  
QY 851 AAATCATGATGCTATCTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 910  
Db 553 AAATCATGATGCTATCTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 612  
QY 911 ATTGCTTACAGAAAGGCTCTTAAAGCCGAGGAGGCGGAGGAGGCTTCACTTCTGCTGCTGCT 970  
Db 613 ATTGCTTACAGAAAGGCTCTTAAAGCCGAGGAGGCGGAGGAGGCTTCACTTCTGCTGCTGCT 672  
QY 971 TTGCTATCCCTCTGAGAAAGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030

Db 673 TTGCTATCCCTTCAGAGACCAAGGAGAACTCTGTGTACCTGTGGAGGAATA-----AT 726  
QY 1031 CGAGTGTGACATGAGTGGCTGAAACCTCAGGAGACTGGACATCCCATGTTGAGCAATG 1090  
Db 727 GTGGTGTGACTGAGGTGATCTG-AGTCTGAGGAGCTGGATATCCCGAGTTCAAGTATG 785  
QY 1091 TCAATGGCATGAGGAGGCGCCCAAGGGCCGCTTCCTTCATGATCCAT 1146  
Db 786 CCAGCAATATCAGGAAGTGCCTCCAGGTGTCCCAACACATCCATCTTTCTATTTCAT 841

## RESULT 3

US-10-029-191-1  
; Sequence 1, Application US/10029191  
; Publication NO. US20020160453A1  
; GENERAL INFORMATION:  
; APPLICANT: CURTIS, RORY A.J.  
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 210147.00XX/5U1  
; CURRENT APPLICATION NUMBER: US/10/029,191  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 09/569,978  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/134,198  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3108  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-10-029-191-1

Query Match 44.4%; Score 559.6; DB 13; Length 3108;  
Best Local Similarity 84.4%; Pred. No. 4e-160;  
Matches 655; Conservative 0; Mismatches 114; Indels 7; Gaps 2;  
QY 371 AGAAGATGCTGCTTCAATAGATGTTTCCCTCGGCTTCTCTGCTTATCTACTGCGG 430  
Db 73 ABAAGATGCTGCTTCAACAGATGCTTCCCTAGCTTCTCTAGTCTCTACTCTGCGG 132  
QY 431 TCAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGAGGCGCTGAGGGCA 490  
Db 133 TCAGAGTCTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGAGGCGCTGAGGGCA 192  
QY 491 ACCCATGAGTGTGCTGCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 550  
Db 193 ATCCCATGAGTGTGAGTGTGATCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAG 252  
QY 551 TGGTGAATGTTCTACAGGCGCGAGGCGGTTAAGATTTCTTATTTACAGTATCGGA 610  
Db 253 TGGTGAATGTTCTACAGGCGCGAGGCGGTTAAGATTTCTTATTTATATGATATCGGA 312  
QY 611 ATGGCCACAGAGAGTGGAGAGCGCTTTCAGGGCGCGCTCAGTGAATGGCAGCAAG 670  
Db 313 ATGGCCACAGAGAGTGGAGAGCGCTTTCAGGGCGCGCTCAGTGAATGGCAGCAAG 372  
QY 671 ACCTGAGAGAGTGTCCATCACTGTGCTCAACGTCATCTGAACGACTCTGGCTTACA 730  
Db 373 ACCTGAGAGAGTGTCCATCACTGTGCTCAACGTCATCTGAACGACTCTGGCTTACA 432  
QY 731 CCTGCAATGTGTCGGGAGTTGAGTTTGAAGGCGCATCGGCCCTTTGTGAGAGAGCG 790  
Db 433 CATGCAATGTGTCGGGAGTTGAGTTTGAAGGCGCATCGGCCCTTTGTGAGAGAGCG 492  
QY 791 GGCTGATCCCTTAAAGAGTCAACAGGAGGCTGGAGAGAGTTCACCTCTGTGTCTCAG 850  
Db 493 GACTGATACCTTTGCGAGTCACTGAAGAGGCGGAGAGAGTTCACCTCTGTGTCTCAG 552  
QY 851 AATCATGATGTACATCTTCTGCTCTTCTCTCACTGCTGCTGCTGCTGCTGCTGCT 910

Db 553 AATCATGATGTACATCTCTCTGCTTCTCCTCACCTTGTGCTGTTTATGAGATGATCT 612  
QY 911 ATTGCTACAGAAAGTCTCAAAAGCCGAGGAGGAGCCCAAGAAACCGCTTCACTACC 970  
Db 613 ATTGCTACAGAAAGTCTCAAAAGCCGAGGAGGAGCCCAAGAAACCGCTTCACTACC 672  
QY 971 TTGCCATCCCATCTCAGAACAGGAGACTCTGGGTACCACTGAGGAGATGAGACAGGA 1030  
Db 673 TTGCTATCCCTTCAGAGAACAGGAGAACTCTGTGTGTACCTGTGGAGGAATA-----AT 726  
QY 1031 GCAGTGTGACATGAGTGGCTGAAACCTCAGGAGACTGGACATCCCATGTTCAAGCAATG 1090  
Db 727 GTGGTGTGACTGAGGTGATCTG-AGTCTGAGGAGCTGGATATCCCGAGTTCAAGTATG 785  
QY 1091 TCAATGGCATGAGGAGGCGCCCAAGGGCCGCTTCCTTCATGATCCAT 1146  
Db 786 CCAGCAATATCAGGAAGTGCCTCCAGGTGTCCCAACACATCCATCTTTCTATTTCAT 841

## RESULT 4

US-09-997-579-3  
; Sequence 3, Application US/09997579  
; Patent No. US20020113203A1  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge University Technical Services  
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated s  
; TITLE OF INVENTION: channel  
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses  
; FILE REFERENCE: 674558-2001  
; CURRENT APPLICATION NUMBER: US/09/997,579  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: PCT/EP00/01783  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60,129,473  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 2220  
; TYPE: DNA  
; ORGANISM: rat  
US-09-997-579-3

Query Match 42.8%; Score 540; DB 10; Length 2220;  
Best Local Similarity 89.3%; Pred. No. 3.4e-154;  
Matches 582; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 371 AGAAGATGCTGCTTCAATAGATGTTTCCCTCGGCTTCTCTGCTTATCTACTGCGG 430  
Db 358 ABAAGATGCTGCTTCAACAGATGCTTCCCTAGCTTCTCTAGTCTCTACTACTGCGG 417  
QY 431 TCAGTGTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGAGGCGCTGAGGGCA 490  
Db 418 TCAGAGTCTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGAGGCGCTGAGGGCA 477  
QY 491 ACCCATGAGTGTGCTGCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 550  
Db 478 ATCCCATGAGTGTGAGTGTGATCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAG 537  
QY 551 TGGTGAATGTTCTACAGGCGCGAGGCGGTTAAGATTTCTTATTTACAGTATCGGA 610  
Db 538 TGGTGAATGTTCTACAGGCGCGAGGCGGTTAAGATTTCTTATATATGATATCGGA 597  
QY 611 ATGGCCACAGAGAGTGGAGAGCGCTTTCAGGGCGCGCTTCAAGTGAATGGCAGCAAG 670  
Db 598 ATGGCCACAGAGAGTGGAGAGCGCTTTCAGGGCGCGCTTCAAGTGAATGGCAGCAAG 657  
QY 671 ACCTGAGAGAGTGTCCATCACTGTGCTCAACGTCATCTCTGAACGACTCTGGCTTACA 730  
Db 658 ACCTGAGAGAGTGTCCATCACTGTGCTCAACGTCATCTCTGAACGACTCTGGCTTACA 717  
QY 731 CCTGCAATGTGTCTCGGAGTTTGAAGGCGCATCGGCCCTTTGTGAGAGAGAGCG 790





```
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEWICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16214
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: Q07699, EVALUE 5.00e-42
; OTHER INFORMATION: NT HIT: U12191.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF953777.1, EVALUE 1.00e-129
US-10-029-386-16214
```

```
Query Match 3.9%; Score 49; DB 12; Length 243;
Best Local Similarity 60.1%; Pred. No. 0.00014;
Matches 104; Conservative 0; Mismatches 60; Indels 9; Gaps 1;

Qy 598 TACGAGTATCGGAATGCGCCACAGGAGGTGAGAGCCCTTTCAGGGCGCGCTGCAGTGG 657
Db 4 TATGAGAATGAGGTGTCAGCTGAGAGGAGATGAGCGCTTCGAGGGCGCGGTGTGG 63

Qy 658 AATGGCAG-----CAAGGACCTGCAGACGTGTCATCACTGTCTCAACGTCACT 708
Db 64 AATGGCAGCGGGCGGACCAAGAGCTGCAGGATCTGTCTATCTTCATCACCAGTCAAC 123

Qy 709 CTGAACGACTCTGGCTCTACACCTGCAATGTGTCGGGAGTTTGAGTTGA 761
Db 124 TACAACCACTCGGGCGGACTAGAGTGCCACGTCTACCGCTCTCTCTTCGGA 176
```

```
RESULT 10
US-10-029-386-2514
; Sequence 2514, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEWICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2514
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
```

```
; OTHER INFORMATION: SWISSPROT HIT: Q07699, EVALUE 6.00e-42
; OTHER INFORMATION: NT HIT: U12191.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: B1754689.1, EVALUE 0.00e+00
US-10-029-386-2514
```

```
Query Match 3.9%; Score 49; DB 12; Length 569;
Best Local Similarity 60.1%; Pred. No. 0.00019;
Matches 104; Conservative 0; Mismatches 60; Indels 9; Gaps 1;

Qy 598 TACGAGTATCGGAATGCGCCACAGGAGGTGAGAGCCCTTTCAGGGCGCGCTGCAGTGG 657
Db 62 TATGAGAATGAGGTGTCAGCTGAGAGGAGATGAGCGCTTCGAGGGCGCGGTGTGG 121

Qy 658 AATGGCAG-----CAAGGACCTGCAGACGTGTCATCACTGTCTCAACGTCACT 708
Db 122 AATGGCAGCGGGCGGACCAAGAGCTGCAGGATCTGTCTATCTTCATCACCAGTCAAC 181

Qy 709 CTGAACGACTCTGGCTCTACACCTGCAATGTGTCGGGAGTTTGAGTTGA 761
Db 182 TACAACCACTCGGGCGGACTAGAGTGCCACGTCTACCGCTCTCTCTTCGGA 234
```

```
RESULT 11
US-10-017-161-1435
; Sequence 1435, Application US/10017161
; Publication No. US20030143688A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: Patenscin ver. 2.1
; SEQ ID NO 1435
; LENGTH: 922
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(922)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(722)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(26)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (62)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (124)..(125)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (127)..(128)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (133)..(134)
; OTHER INFORMATION: a, t, c, g, unknown or other
```

```
FEATURE:
NAME/KEY: modified_base
LOCATION: (138)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (142)..(145)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (147)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (151)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (311)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (322)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (326)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (345)..(349)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (351)..(361)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (364)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (366)..(368)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (370)..(380)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (384)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (386)..(388)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (392)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (394)..(396)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (403)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (413)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
```

```
NAME/KEY: modified_base
LOCATION: (415)..(417)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (422)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (438)..(439)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (441)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (443)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (445)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (465)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (469)..(471)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (487)..(488)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (490)..(491)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (495)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (497)..(498)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (501)..(505)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (519)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (529)..(530)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (540)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (548)..(549)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (564)..(568)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
```

/ LOCATION: (571)  
/ OTHER INFORMATION: a, t, c, g, unknown or other  
/ FEATURE:  
/ NAME/KEY: modified\_base  
/ LOCATION: (573)

Query Match 3.3%; Score 41.2; DB 12; Length 922;

Best Local Similarity 40.9%; Pred. No. 0.056;  
Matches 70; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 12 GAGCTGAGCTTACCTTGGCGCAACAGGAGGAGGCGGCGGAGCTGGAGTT 71  
DB 363 GNGNNGNN 422  
QY 72 CCGGGGTTGGCGGAGGAGGAGCTGCTCCGTGTGTGCTGAGCGCGGAGAGCGGCGGA 131  
DB 423 GCGGGGGGGGGGNGNGNGNGGCGGGGGGGGGGGGNGGNNNGGGGGGG 482  
QY 132 GCGGCTGATCGCTCCCTCGAACTGGGAGGTCAGTGGGTGCTTAGGG 182  
DB 483 GCGGNNNGGNGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 533

## RESULT 12

US-09-771-161A-58  
/ Sequence 58, Application US/09771161A

/ Patent No. US20020110811A1

/ GENERAL INFORMATION:

/ APPLICANT: LEVINE, et al.

/ TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES

/ FILE REFERENCE: 802620-2005.1

/ CURRENT APPLICATION NUMBER: US/09/771,161A

/ CURRENT FILING DATE: 2001-01-26

/ PRIOR APPLICATION NUMBER: 09/724,676

/ PRIOR FILING DATE: 2000-11-28

/ PRIOR APPLICATION NUMBER: 136776

/ PRIOR FILING DATE: 2000-06-15

/ PRIOR APPLICATION NUMBER: 135619

/ PRIOR FILING DATE: 2000-04-12

/ NUMBER OF SEQ ID NOS: 273

/ SOFTWARE: Patent in version 3.0

/ SEQ ID NO 58

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-09-771-161A-58

Query Match 3.3%; Score 41; DB 10; Length 1097;

Best Local Similarity 52.0%; Pred. No. 0.069; Indels 85; Gaps 0;  
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1 CCCTCCCTTCCGAGCTGAGCTTACCTTGGCGCAACAGGAGGAGGCGGAGTGG 60  
DB 23 CCCCCGCTCGGGGCTGTGAGCGGCTCGGGGCGGGGCTGGCGCGGCGGCGG 82  
QY 61 AAGCTGAGTTCCGGGTTGGCGGAGGAGGAGCTGCTCCGTGTGTGAGCGCGGAGA 120  
DB 83 CCGACGCTCTCTTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 142  
QY 121 GCGGGCGGAGCGGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGTGCT 177  
DB 143 GCGGGCGGAGTTGCTTCCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCT 199

## RESULT 13

US-09-771-161A-59

/ Sequence 59, Application US/09771161A

/ Patent No. US20020110811A1

/ GENERAL INFORMATION:

/ APPLICANT: LEVINE, et al.

/ TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES

/ FILE REFERENCE: 802620-2005.1

/ CURRENT APPLICATION NUMBER: US/09/771,161A

/ CURRENT FILING DATE: 2001-01-26  
/ PRIOR APPLICATION NUMBER: 09/724,676  
/ PRIOR FILING DATE: 2000-11-28  
/ PRIOR APPLICATION NUMBER: 136776  
/ PRIOR FILING DATE: 2000-06-15  
/ PRIOR APPLICATION NUMBER: 135619  
/ PRIOR FILING DATE: 2000-04-12  
/ NUMBER OF SEQ ID NOS: 273  
/ SOFTWARE: Patent in version 3.0  
/ SEQ ID NO 59  
/ LENGTH: 3382  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: -  
/ LOCATION: (1)...(3382)  
/ OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'  
US-09-771-161A-59

Query Match 3.3%; Score 41; DB 10; Length 3382;

Best Local Similarity 52.0%; Pred. No. 0.1;  
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1 CCCTCCCTTCCGAGCTGAGCTTACCTTGGCGCAACAGGAGGAGGCGGCGGAGTGG 60  
DB 23 CCCCCGCTCGGGGCTGTGAGCGGCTCGGGGCGGGGCTGGCGCGGCGGCGG 82  
QY 61 AAGCTGAGTTCCGGGTTGGCGGAGGAGGAGCTGCTCCGTGTGTGAGCGCGGAGA 120  
DB 83 CCGACGCTCTCTTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 142  
QY 121 GCGGGCGGAGCGGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGTGCT 177  
DB 143 GCGGGCGGAGTTGCTTCCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCT 199

## RESULT 14

US-09-291-417-102

/ Sequence 102, Application US/09291417A

/ Publication No. US20030050230A1

/ GENERAL INFORMATION:

/ APPLICANT: FLOWMAN, GREGORY

/ APPLICANT: MARTINEZ, RICARDO

/ APPLICANT: WHYTE, DAVID

/ TITLE OF INVENTION: STR20-RELATED PROTEIN KINASES

/ FILE REFERENCE: 240/300

/ CURRENT APPLICATION NUMBER: US/09/291,417A

/ CURRENT FILING DATE: 1999-04-13

/ EARLIER APPLICATION NUMBER: US 60/081,784

/ EARLIER FILING DATE: 1998-04-14

/ NUMBER OF SEQ ID NOS: 147

/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO 102

/ LENGTH: 2806

/ TYPE: DNA

/ ORGANISM: Full Length Mammalian (Human) PKA5

US-09-291-417-102

Query Match

Best Local Similarity 49.8%; Pred. No. 0.17;

Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 23 ACCCTGGGCGCAACAGGAGGAGGCGGCGGAGTGGAGTGGAGTTCGGGGTGGGC 82  
DB 567 ACCACGGCCAGAGGGGGGCGCCAGGGAAGGAGGAGGAGGCGGCGGCGGCGG 626  
QY 83 GGGGAGGCGAGCTGTCCGTGCTGCTGAGCGCCCGGAGAGCGGCGGCGGCGGCGG 142  
DB 627 GAGCGGGGTGGCGGCGAGTGTGTACAGGCGAGCGGCGGCGGCGGCGGCGGCGG 586  
QY 143 GCTCCCTCGAAGTGGGGAGGTCGAGTGGGGTGGCTTTAGGCGCCCAAGCCCCCGGCT 202  
DB 687 TCAGGGAGGAGGCTCAGGGGGTCCCCCAGGAGTCTCTCCCGGAGCAAAACGCGCCCT 746



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:16:11, Search time 2705.05 Seconds  
(without alignments)  
9799.969 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_376\_1023

Perfect score: 648

Sequence: 1 atgctgccttcaatagatt.....cggtagcagtgagggaatag 648

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl:

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sv.\*

13: gb\_un.\*

14: gb\_vt.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_on.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vt.\*

30: em\_hg\_hum.\*

31: em\_hg\_inv.\*

32: em\_hg\_other.\*

33: em\_hg\_mus.\*

34: em\_hg\_pln.\*

35: em\_hg\_rod.\*

36: em\_hg\_mam.\*

37: em\_hg\_vrt.\*

38: em\_sy.\*

39: em\_hgo\_hum.\*

40: em\_hgo\_mus.\*

41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	648	100.0	1261	6	AX039100	Sequence
2	648	100.0	1261	9	AJ243396	Homo sapi
3	646.4	99.8	4052	9	HSMB01563	Homo sapi
4	646.4	99.8	5306	9	AB0322984	Homo sapi
5	627.2	96.8	3296	9	AB097521	Macaca fa
6	537.2	82.9	670	10	AY049036	Mus muscu
7	536.6	82.8	2220	6	AX039099	Sequence
8	536.6	82.8	2220	10	RNO243395	Rattus no
9	536.6	82.8	2632	6	AX048005	Sequence
10	536.6	82.8	3107	10	AF378093	Rattus no
11	536.6	82.8	3108	6	AX047984	Sequence
12	534.6	82.5	645	6	AX048004	Sequence
13	422	65.1	471	6	BD059018	Secreted
14	229.6	35.4	127347	2	AP000682	Homo sapi
15	229.6	35.4	172546	2	AC024604	Homo sapi
16	229.6	35.4	178169	9	AP002765	Homo sapi
17	229.6	35.4	178431	2	AP002749	Homo sapi
18	229.6	35.4	181471	9	AC069539	Homo sapi
19	228	35.2	144833	2	AC021981	Homo sapi
20	228	35.2	149800	2	AC021981	Homo sapi
21	179	27.6	143624	2	AC135353	Mus muscu
22	174.2	26.9	214306	2	AC128723	Rattus no
23	167.8	25.9	144833	2	AC063921	Homo sapi
24	148.2	22.9	657	4	OC035382	Oryctolagus
25	148.2	22.9	657	6	AX048006	Sequence
26	145.4	22.4	1404	9	HUMSCN18A	Human sodiu
27	143.8	22.2	1335	9	HUMVSGC1B	Homo sapien
28	138.6	21.4	1125	10	MMU46681	Mus musculu
29	138.6	21.4	1340	10	MMU85786	Mus musculu
30	138.6	21.4	1490	6	AX401378	Sequence
31	138.6	21.4	1490	10	RAT01378	Rattus norv
32	138.6	21.4	1557	10	BC039140	Mus muscu
33	138.6	21.4	1568	10	BC009652	Mus muscu
34	119.6	18.5	1248	9	BC021266	Homo sapi
35	92.6	14.3	74467	2	AC013674	Homo sapi
36	81.4	12.6	807	6	AX354520	Sequence
37	81.4	12.6	974	6	AX354521	Sequence
38	74.4	11.5	136892	5	AL954724	Zebrafish
39	73	11.3	850	6	AX105630	Sequence
40	73	11.3	850	6	AX111728	Sequence
41	73	11.3	850	10	AF182949	Rattus no
42	64	9.9	178431	2	AP002749	Homo sapi
43	55	8.5	66980	2	AC013796	Homo sapi
44	55	8.5	66980	2	AC013796	Homo sapi
45	55	8.5	149800	2	AC021981	Homo sapi

## ALIGNMENTS

### RESULT 1

AX039100

LOCUS

DEFINITION

AX039100

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1

AX039100

LOCUS

DEFINITION

AX039100

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AX039100 Sequence 4 from Patent WO0063367. 1261 bp DNA linear PAT 18-NOV-2000

Homo sapiens (human)  
Homo sapiens  
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

Cox, P., Dixon, A., Jackson, A. and Morgan, K.  
A novel family of beta sub-unit proteins from a voltage-gated sodi  
um channel, nucleic acids encoding them and therapeutic or

diagnostic uses there of  
Patent: WO 0063367-A 4 26-OCT-2000;  
WARNER-LAMBERT COMPANY (US); Cambridge University Technical  
Services Limited (GB)  
FEATURES  
Location/Qualifiers  
1..1261  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 263 a 366 c 372 g 260 t  
ORIGIN  
Query Match 100.0%; Score 648; DB 6; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 2.7e-142;  
Matches 648; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCTGCTTCAATAGATTGTTTCCCTGCTTCTCTCGTGTATCTACTGCTGCTCAGT 60  
Db |||||||  
376 ATGCTGCTTCAATAGATTGTTTCCCTGCTTCTCTCGTGTATCTACTGCTGCTCAGT 435  
Qy 61 GTCTGCTTCTCTGTGTGTGAAGTGCCTTCGAGACGAGGCGGTCGAGGCAACCCC 120  
Db |||||||  
436 GTCTGCTTCTCTGTGTGTGAAGTGCCTTCGAGACGAGGCGGTCGAGGCAACCCC 495  
Qy 121 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Db |||||||  
496 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555  
Qy 181 GAATGGTCTTACAGCCCGGAGGCGGTAAGATTTCCTTATTTACGATATCGGAATGGC 240  
Db |||||||  
556 GAATGGTCTTACAGCCCGGAGGCGGTAAGATTTCCTTATTTACGATATCGGAATGGC 615  
Qy 241 CACGAGAGTGGAGAGGCGGCTTTCAGGGGCGGCTGAGTGGATGCGAGGAGGAGCTG 300  
Db |||||||  
616 CACGAGAGTGGAGAGGCGGCTTTCAGGGGCGGCTGAGTGGATGCGAGGAGGAGCTG 675  
Qy 301 CAGGAGCTGCTCATCAGTGTCTCAACCTCACTCTGAACGACTCTGGGCTCTACACCTGC 360  
Db |||||||  
676 CAGGAGCTGCTCATCAGTGTCTCAACCTCACTCTGAACGACTCTGGGCTCTACACCTGC 735  
Qy 361 AATGTGCTCCGGAGTTTGTAGTTGAGCGGCATCGCCCTTTGTGAAGACGACGCGGCTG 420  
Db |||||||  
736 AATGTGCTCCGGAGTTTGTAGTTGAGCGGCATCGCCCTTTGTGAAGACGACGCGGCTG 795  
Qy 421 ATCCCTCTAAGAGTCAACGAGAGGCTGAGGAGCTTCACTCTGTGCTCTGAGAAATC 480  
Db |||||||  
796 ATCCCTCTAAGAGTCAACGAGAGGCTGAGGAGCTTCACTCTGTGCTCTGAGAAATC 855  
Qy 481 ATGATGTACATCTTCTGCTTCTCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db |||||||  
856 ATGATGTACATCTTCTGCTTCTCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 915  
Qy 541 TACGAAAGGCTTCAAAAGCGAGAGGAGGCGGCAAGAAACGCGTCTGACTACCTTGCC 600  
Db |||||||  
916 TACGAAAGGCTTCAAAAGCGAGAGGAGGCGGCAAGAAACGCGTCTGACTACCTTGCC 975  
Qy 601 ATCCCTCTGAGAACAGAGAGTCTCGGTTACAGTGGAGGATAG 648  
Db |||||||  
976 ATCCCTCTGAGAACAGAGAGTCTCGGTTACAGTGGAGGATAG 1023  
RESULT 2  
HSA243396  
LOCUS  
DEFINITION Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit  
(scn3b gene).  
ACCESSION AJ243396  
VERSION AJ243396.2  
KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K.,  
Pinnock, R.D., Hughes, J., Richardson, P.J., Mizuguchi, K. and  
Jackson, A.P.  
TITLE beta 3: an additional auxiliary subunit of the voltage-sensitive  
sodium channel that modulates channel gating with distinct kinetics  
Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)  
JOURNAL 20160948  
MEDLINE 1068874  
PUBMED  
REFERENCE 2 Morgan, K.  
TITLE Direct Submission  
JOURNAL Submitted (28-JUN-1999) Morgan K., Biochemistry, University of  
Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
REMARK Revised by [4]  
REFERENCE 3 (bases 1 to 1261)  
Morgan, K.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAR-2000) Morgan K., Biochemistry, University of  
Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
COMMENT On Mar 14, 2000 this sequence version replaced gi:7160974.  
FEATURES  
Location/Qualifiers  
1..1261  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="brain"  
1..1261  
/genes="scn3b"  
1..375  
/genes="scn3b"  
376..1024  
/genes="scn3b"  
/function="modulates channel gating kinetics"  
/codon\_start=1  
/evidence="experimental"  
/product="voltage-gated sodium channel beta-3 subunit"  
/protein\_id="CA576825.1"  
/db\_xref="GI:7160975"  
/translation="MPANRLPLASLVLYVSVCFVCPVSPSTEAVQGNPKLR  
CISCKREEVATVYVYRFGKGLIYRNGHQEVESPFQRLLONGSKLDQ  
VSITLVNLTSLYTCNVRFEPFHAHPFKTLLIPLRVTEAGEDTSSVSEI  
MYVLLVFLTLWLLIEMIVCYRKVSKAEAAQENASDYLAIPSENKENSAPVVEE"  
1025..->1261  
/genes="scn3b"  
BASE COUNT 263 a 366 c 372 g 260 t  
ORIGIN  
Query Match 100.0%; Score 648; DB 9; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 2.7e-142;  
Matches 648; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCTGCTTCAATAGATTGTTTCCCTGCTTCTCTCGTGTATCTACTGCTGCTCAGT 60  
Db |||||||  
376 ATGCTGCTTCAATAGATTGTTTCCCTGCTTCTCTCGTGTATCTACTGCTGCTCAGT 435  
Qy 61 GTCTGCTTCTCTGTGTGTGAAGTGCCTTCGAGACGAGGCGGTCGAGGCAACCCC 120  
Db |||||||  
436 GTCTGCTTCTCTGTGTGTGAAGTGCCTTCGAGACGAGGCGGTCGAGGCAACCCC 495  
Qy 121 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Db |||||||  
496 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555  
Qy 181 GAATGGTCTTACAGCCCGGAGGCGGTAAGATTTCCTTATTTACGATATCGGAATGGC 240  
Db |||||||  
556 GAATGGTCTTACAGCCCGGAGGCGGTAAGATTTCCTTATTTACGATATCGGAATGGC 615  
Qy 241 CACGAGAGTGGAGAGGCGGCTTTCAGGGGCGGCTGAGTGGATGCGAGGAGGAGCTG 300  
Db |||||||  
616 CACGAGAGTGGAGAGGCGGCTTTCAGGGGCGGCTGAGTGGATGCGAGGAGGAGCTG 675  
Qy 301 CAGGAGCTGCTCATCAGTGTCTCAACCTCACTCTGAACGACTCTGGGCTCTACACCTGC 360

676 CAGGACGTTGTCATCACTGTGCTCAACCTCACTCTGAACGACTCTGGGCTCTACACCTGC 735  
361 AATGTGTCCTCCGGAGTTGAGTTTGAGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 420  
736 AATGTGTCCTCCGGAGTTGAGTTTGAGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 795  
421 ATCCCCCTAAGAGTCACCGAGGAGCTGGAGAGACTTCACTCTGTGGTCTCAGAAATC 480  
796 ATCCCCCTAAGAGTCACCGAGGAGCTGGAGAGACTTCACTCTGTGGTCTCAGAAATC 855  
481 ATGATGATACCTCTTCTGCTTCTTCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
856 ATGATGATACCTCTTCTGCTTCTTCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 915  
541 TACAGAAAGGCTCTCAAAAGCCGAGAGGACGACCAAGAAACGGCTCTGACTACCTTGC 600  
916 TACAGAAAGGCTCTCAAAAGCCGAGAGGACGACCAAGAAACGGCTCTGACTACCTTGC 975  
601 ATCCCATCTGAGAACAGGAACTCTCGGTACAGTGGAGGAATAG 648  
976 ATCCCATCTGAGAACAGGAACTCTCGGTACAGTGGAGGAATAG 1023

RESULT 3  
HSM801563  
LOCUS  
DEFINITION Homo sapiens mRNA; cDNA DKFZp761F182 (from clone DKFZp761F182); complete cds.  
VERSION AL136589  
KEYWORDS AL136589.1 GI:13276680  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS Homo sapiens  
JOURNAL Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
COMMENT 1 (bases 1 to 4052)  
Ottewaelder,B., Obermaier,B., Mewes,H.W., Weil,B. and Wiemann,S.  
Submitted (12-MAR-2002) MIPS, Am Klopferstutz 18a, D-82152 Martinsried, GERMANY  
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by MedGenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761F182) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

## FEATURES

Location/Qualifiers  
1..4052  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp761F182"  
/tissue\_type="amygdala"  
/clone\_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
1..4052  
/gene="DKFZp761F182"  
804..1451  
/gene="DKFZp761F182"  
/note="scn3b (Homo sapiens)"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAB6524.1"  
/db\_xref="GI:13276681"  
/translaton="MPAFNLFPLASLIYWSVCFPPVEPSPSEAVQGNPKLR  
CISCMKEEVEATVWVFYRPEGKDFLIYVRNGHOEVSPQGRLOWGSKDLQD  
VSITVNLVNDLSGLYTCNVSREPEFPAHRPFVKTLLIPLRYTEBAGEFTSVSEI

## gene

## CDS

polya\_site  
BASE COUNT 994 a 1033 c 1036 g 989 t  
ORIGIN  
Query Match 99.8%; Score 646.4; DB 9; Length 4052;  
Best Local Similarity 99.8%; Pred. No. 6.1e-142;  
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGCTGCTCCCTCAATAGATTGTTCCCTGGCTTCTCTCGTCTATCTACTGGTCACT 60  
Db 804 ATGCTGCTCCCTCAATAGATTGTTCCCTGGCTTCTCTCGTCTATCTACTGGTCACT 863  
Qy 61 GTCTGCTCCCTGTTGTGGAGTCCCTCCGAGACGAGGCGCTGAGGCGCAACCCC 120  
Db 864 GTCTGCTCCCTGTTGTGGAGTCCCTCCGAGACGAGGCGCTGAGGCGCAACCCC 923  
Qy 121 ATGAAGCTGCGCTGCATCTCTCTCATGAAGAGAGAGAGTGGAGGCCACACGGTGGT 180  
Db 924 ATGAAGCTGCGCTGCATCTCTCTCATGAAGAGAGAGAGTGGAGGCCACACGGTGGT 983  
Qy 181 GAATGTTCTACAGGCCCGAGGCGCGTAAAGATTCTTATTACGATATCGAATGCG 240  
Db 984 GAATGTTCTACAGGCCCGAGGCGCGTAAAGATTCTTATTACGATATCGAATGCG 1043  
Qy 241 CACGAGGAGTGGAGAGCCCTTTTTCAGGGGCGCTGAGTGGATGCGCAGGACCTG 300  
Db 1044 CACGAGGAGTGGAGAGCCCTTTTTCAGGGGCGCTGAGTGGATGCGCAGGACCTG 1103  
Qy 301 CAGGACGTGTCCATCTGCTCAACGTCACTCTGAACGACTCTGGCTCTACACCTGC 360  
Db 1104 CAGGACGTGTCCATCTGCTCAACGTCACTCTGAACGACTCTGGCTCTACACCTGC 1163  
Qy 361 AATGTGTCCTCCGGAGTTGAGTTTGAGGCGCATCGGCCCTTCTGAGAGACGCGCTG 420  
Db 1164 AATGTGTCCTCCGGAGTTGAGTTTGAGGCGCATCGGCCCTTCTGAGAGACGCGCTG 1223  
Qy 421 ATCCCTTAAGAGTCAACGAGGAGGCTGGAGAGACTTCACTCTGTGGTCTCAGAAATC 480  
Db 1224 ATCCCTTAAGAGTCAACGAGGAGGCTGGAGAGACTTCACTCTGTGGTCTCAGAAATC 1283  
Qy 481 ATGATGATACCTCTTCTGCTTCTCACTCCCTGTGCTGCTATCGAGATGATATATGC 540  
Db 1284 ATGATGATACCTCTTCTGCTTCTTCTCCACCTTGTGGCTGCTCATCGAGATGATATATGC 1343  
Qy 541 TACAGAAAGTCTCAAAAGCCGAGGAGGCGCCCAAGAAACGGCTCTGACTACCTTGC 600  
Db 1344 TACAGAAAGTCTCAAAAGCCGAGGAGGCGCCCAAGAAACGGCTCTGACTACCTTGC 1403  
Qy 601 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACAGTGGAGGAATAG 648  
Db 1404 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACAGTGGAGGAATAG 1451

## RESULT 4

LOCUS AB032984  
DEFINITION Homo sapiens mRNA for KIAA1158 protein, partial cds.  
ACCESSION AB032984  
VERSION AB032984.1 GI:6330135  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
Hirose,M., Nagase,T., Ishikawa,K., Kikuno,R., Nomura,N. and Ohara,O.  
TITLE Characterization of cDNA clones selected by the GeneMark analysis from size-fractionated cDNA libraries from human brain  
JOURNAL DNA Res. 6 (5), 329-336 (1999)  
MEDLINE 20039618

```
10574461
2 (bases 1 to 5306)
REFERENCE
  Chakraborty, O., Nagase, T. and Kikuno, R.
  Direct Submission
  Submitted (04-DEC-1999) Osamu Ohara, Kazusa DNA Research Institute,
  Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
  292-0812, Japan (E-mail:cdna@kazusa.or.jp)
  URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
  Fax: +81-438-52-3914
  Location/Qualifiers
    1..5306
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="h100081"
      /sex="male"
      /tissue_type="brain"
      /clone_lib="pBluescriptII SK plus"
      /dev_stage="adult"
    1..5306
      /gene="KIAA1158"
      /note="start codon is not identified."
      /codon_start=3
      /product="KIAA1158 protein"
      /protein_id="BA86472.1"
      /db_xref="GI:6330336"
      /translation="RWSGSESLGSRQPKMPAFNRLFLPLASLVLIYWSVFCVVEV
      PSEAVQNPMPKRCICSMKREVEATVVEWYFPEGKDFLYEYRNGHVEVSP
      FQGRLOWGSKDQVSTVLNLTNSGLYTCNVSRFEFEHPPFKVTRLLPLRV
      TEEAGEFTSVVSEIMVYLLVFLTLWLLIEMIYCRKVSKEAEQAENASDYLAI
      ENKENSAPVBER"
BASE COUNT 1375 a 1311 c 1111 g 1509 t
ORIGIN
Query Match 99.8%; Score 646.4; DB 9; Length 5306;
Best Local Similarity 99.8%; Pred. No. 6e-142;
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 60
Db 1 ATGCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 107
QY 61 GTCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 120
Db 108 GTCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 167
QY 121 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 168 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 227
QY 181 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 228 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
QY 241 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 288 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 347
QY 301 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 348 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
QY 361 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 408 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
QY 421 ATCCCTCTAAGAGTACCGAGAGGCTCGAGAGGACTTCACTCTGCTGCTGCTGCTGCTGCT 480
Db 468 ATCCCTCTAAGAGTACCGAGAGGCTCGAGAGGACTTCACTCTGCTGCTGCTGCTGCTGCT 527
QY 481 ATGATGTACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
```

```
10574461
2 (bases 1 to 5306)
REFERENCE
  Chakraborty, O., Nagase, T. and Kikuno, R.
  Direct Submission
  Submitted (04-DEC-1999) Osamu Ohara, Kazusa DNA Research Institute,
  Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
  292-0812, Japan (E-mail:cdna@kazusa.or.jp)
  URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
  Fax: +81-438-52-3914
  Location/Qualifiers
    1..5306
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="h100081"
      /sex="male"
      /tissue_type="brain"
      /clone_lib="pBluescriptII SK plus"
      /dev_stage="adult"
    1..5306
      /gene="KIAA1158"
      /note="start codon is not identified."
      /codon_start=3
      /product="KIAA1158 protein"
      /protein_id="BA86472.1"
      /db_xref="GI:6330336"
      /translation="RWSGSESLGSRQPKMPAFNRLFLPLASLVLIYWSVFCVVEV
      PSEAVQNPMPKRCICSMKREVEATVVEWYFPEGKDFLYEYRNGHVEVSP
      FQGRLOWGSKDQVSTVLNLTNSGLYTCNVSRFEFEHPPFKVTRLLPLRV
      TEEAGEFTSVVSEIMVYLLVFLTLWLLIEMIYCRKVSKEAEQAENASDYLAI
      ENKENSAPVBER"
BASE COUNT 1375 a 1311 c 1111 g 1509 t
ORIGIN
Query Match 99.8%; Score 646.4; DB 9; Length 5306;
Best Local Similarity 99.8%; Pred. No. 6e-142;
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 60
Db 1 ATGCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 107
QY 61 GTCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 120
Db 108 GTCTGCTCCCTCAATAGATTGTTCCCTGCTGCTTCTCTGCTTATCTACTGCTGCTAGT 167
QY 121 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 168 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 227
QY 181 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 228 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
QY 241 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 288 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 347
QY 301 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 348 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
QY 361 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 408 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
QY 421 ATCCCTCTAAGAGTACCGAGAGGCTCGAGAGGACTTCACTCTGCTGCTGCTGCTGCTGCT 480
Db 468 ATCCCTCTAAGAGTACCGAGAGGCTCGAGAGGACTTCACTCTGCTGCTGCTGCTGCTGCT 527
QY 481 ATGATGTACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
```



```

VERSION      AX039099.1  GI:11229275
KEYWORDS     Rattus sp.
SOURCE       Rattus sp.
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE    Cox,P., Dixon,A., Jackson,A. and Morgan,K.
AUTHORS      A novel family of beta sub-unit proteins from a voltage-gated sodi
TITLE        um channel, nucleic acids encoding them and therapeutic or
              diagnostic uses there of
JOURNAL      Patent: WO 0063367-A 3 26-OCT-2000;
              WARNER-LAMBERT COMPANY (US) ; Cambridge University Technical
              Services Limited (GB)
FEATURES     Location/Qualifiers
              source
                1..2220
                /organism="Rattus sp."
                /mol_type="genomic DNA"
                /db_xref="taxon:10118"
BASE COUNT   573 a 557 c 561 g 529 t
ORIGIN
Query Match      82.8%; Score 536.6; DB 6; Length 2220;
Best Local Similarity 89.3%; Pred. No. 5e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 1 ATGCCTGCTTCAATAGATGTTTCCCTGGGTTCTCTCGTCTTATCTACTGGTCACT 60
DB 363 ATGCCTGCTTCAACAGATGTTGCTCCCTAGTCTTCTAGTCTCATCTACTGGTCA 422
QY 61 GTCCTCTCCCTGCTGTGTGAAGTCCCTCGAGAGAGGAGGCGGCGTCCAGGCAACCC 120
DB 423 GTCCTCTCCCTGCTGTGTGAAGTCCCTCGAGAGAGAGGCGGCGTCCAGGCAACCC 482
QY 121 ATGAAGCTGCTGCTGATCTCTCTGATGAAGAGAGAGGAGGAGGAGGAGGAGGAGG 180
DB 483 ATGAAGCTGAGTGCATCTCTGATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
QY 181 GAATGCTTCTACAGCGCGGAGGCGGTAAGATTTCTTATTTACGATGATCGGAATGC 240
DB 543 GAGTGGTCTTACAGCGCTGAGGCGGTAAGATTTCTTATTTACGATGATCGGAATGC 602
QY 241 CACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 603 CACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
QY 301 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 663 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
QY 361 AATGCTGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 723 AATGCTGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
QY 421 ATCCCCCTAAGAGTCAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 783 ATACCTTTGCGAGTCACTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 842
QY 481 ATGATGATACATCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
DB 843 ATGATGATACATCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 902
QY 541 TACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 903 TACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 962
QY 601 ATCCCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 647
DB 963 ATCCCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1009

```

RESULT 8

```

RNO243395      2220 bp  mRNA  linear  ROD 29-JAN-2003
LOCUS          Rattus norvegicus mRNA for voltage-gated sodium channel beta-3
DEFINITION     subunit.
ACCESSION      AJ243395
VERSION        AJ243395.2  GI:7242802
KEYWORDS       scn3b gene; voltage-gated sodium channel beta-3 subunit.
SOURCE         Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
REFERENCE      Morgan,K., Stevens,E.B., Shaw,B., Cox,P., Dixon,A.K., Lee,K.,
AUTHORS        Pinnock,R.D., Hughes,J., Richardson,P.J., Mizuguchi,K. and
                Jackson,A.P.
TITLE          beta 3: an additional auxiliary subunit of the voltage-sensitive
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)
MEDLINE        20160948
PUBMED         10688874
REFERENCE      Morgan,K.
AUTHORS        Direct Submission
TITLE          Submitted (28-JUN-1999) Morgan K., Biochemistry, University of
JOURNAL        Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
REMARK         Revised By [3]
REFERENCE      Morgan,K.
AUTHORS        Direct Submission
TITLE          Submitted (09-MAR-2000) Morgan K., Biochemistry, University of
JOURNAL        Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
COMMENT        On Mar 14, 2000 this sequence version replaced gi:7161888.
FEATURES       Location/Qualifiers
                1..2220
                /organism="Rattus norvegicus"
                /mol_type="mRNA"
                /db_xref="taxon:10116"
                /tissue_type="brain"
                /tissue_lib="lambda ZAP"
                1..2220
                /gene="scn3b"
                <1..362
                /gene="scn3b"
                363..1010
                /gene="scn3b"
                /function="modulates channel opening kinetics"
                /evidence=experimental
                /product="voltage-gated sodium channel beta-3 subunit"
                /protein_id="CA876838.1"
                /db_xref="GI:7161889"
                /db_xref="GOA:Q9UK00"
                /db_xref="SPTREMBL:Q9JK00"
                /translation="MPAFNRLPLASLIYVVRVCPVCPVSEAVQGNPKMLR
                CISMKREVEATVVEWVFRPEGKDFLYEVNRNGHVEFQGLRWNGSKDLQD
                VSTVLNVTINDSLGATCNVSRFEFEHPPVKTFLIPLRVTEAGEDPTSVVSEI
                MMVILLVFLTLHLFIEMTYCYRKVKRAEAAQENASDYLAIIPSENKENSVPVEE"
                1011..>2220
                /gene="scn3b"
BASE COUNT     573 a 557 c 561 g 529 t
ORIGIN
Query Match    82.8%; Score 536.6; DB 10; Length 2220;
Best Local Similarity 89.3%; Pred. No. 5e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 1 ATGCCTGCTTCAATAGATGTTTCCCTGGGTTCTCTCGTCTTATCTACTGGTCACT 60
DB 363 ATGCCTGCTTCAACAGATGTTGCTCCCTAGTCTTCTAGTCTCATCTACTGGTCA 422
QY 61 GTCCTTCCCTGTGTGTGAAGTCCCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120

```

```
Db 423 GTCTGCTTCCCTGTGTGTGGAAAGTGCCCTCGAGAGACAGAGCGGTGCGAGCAATCCC 482
Qy 121 ATGAAGCTGCGCTGCATCTCTCGCATGAAGAGAGAGAGGTGGAGGCCACACCGTGTG 180
Db 483 ATGAAGCTGAGTGCATCTCTCGCATGAAGAGAGAGAGGTGGAGGCCACACCGTGTG 542
Qy 181 GAATGGTTCTACAGCCCGAGGCGGTAAAGATTTCTTATTTACGATATCGGAATGGC 240
Db 543 GAGTGGTTCTACAGCCCTGAGGCGGTAAAGATTTCTTATTTACGATATCGGAATGGC 602
Qy 241 CACAGAGAGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGGAATGCGAGAGACCTG 300
Db 603 CACAGAGAGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGGAATGCGAGAGACCTG 662
Qy 301 CAGAGCGTGTCCATCACTGTCTCAACGTCACCTCTGAACGACTCTGGCCCTCTACACCTGC 360
Db 663 CAGAGCGTGTCCATCACTGTCTCAACGTCACCTCTGAACGACTCTGGCCCTCTACACCTGC 722
Qy 361 AATGTGTCGGGAGTTGAGTTTGGCGCATCGGCCCTTTGTGAAGAGACCGCGCTG 420
Db 723 AATGTGTCGGGAGTTGAGTTTGGCGCATCGGCCCTTTGTGAAGAGACCGCGCTG 782
Qy 421 ATCCCTTTCAGAGTCCAGAGAGGCTGGAGAGACTTCACTCTGTGTCTCAGAAATC 480
Db 783 ATACCTTTCAGAGTCCAGAGAGGCTGGAGAGACTTCACTCTGTGTCTCAGAAATC 842
Qy 481 ATGATGTACATCTTCTGCTTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 843 ATGATGTACATCTTCTGCTTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
Qy 541 TACAGAAAGTCTCAAAAGCGGAGAGAGCGACCCCAAGAAACGCGTCTGACTACCTTGC 600
Db 903 TACAGAAAGTCTCTAAGCGCGAAGAGAGCGACCCCAAGAAATGCGTCTGACTACCTTGC 962
Qy 601 ATCCCTTTCAGAGTCCAGAGAGACTCTGCGGTACCGAGTGGAGGAATA 647
Db 963 ATCCCTTTCAGAGTCCAGAGAGACTCTGCGGTACCGAGTGGAGGAATA 1009

RESULT 9
AX048005
LOCUS
DEFINITION Sequence 22 from Patent WO069912.
ACCESSION AX048005
VERSION AX048005.1 GI:11876883
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITILE
JOURNAL
FEATURES
source
BASE COUNT 693 a 636 c 611 g 692 t
ORIGIN
Query Match 82.8%; Score 536.6; DB 6; Length 2632;
Best Local Similarity 89.3%; Pred. No. 5e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
Qy 1 ATGCTGCTTCCCTGTGTGTGGAAAGTGCCCTCGAGAGAGCGGTGCGAGCAATCCC 60
Db 78 ATGCTGCTTCCCTGTGTGTGGAAAGTGCCCTCGAGAGAGCGGTGCGAGCAATCCC 137
Qy 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCCTCGAGAGAGCGGTGCGAGCAATCCC 120
```

```
Db 138 GTCTGCTTCCCTGTGTGTGGAAAGTGCCCTCGAGAGACAGAGCGGTGCGAGCAATCCC 197
Qy 121 ATGAAGCTGCGCTGCATCTCTCGCATGAAGAGAGAGAGGTGGAGGCCACACCGTGTG 180
Db 198 ATGAAGCTGAGTGCATCTCTCGCATGAAGAGAGAGAGGTGGAGGCCACACCGTGTG 257
Qy 181 GAATGGTTCTACAGCCCGAGGCGGTAAAGATTTCTTATTTACGATATCGGAATGGC 240
Db 258 GAGTGGTTCTACAGCCCTGAGGCGGTAAAGATTTCTTATTTACGATATCGGAATGGC 317
Qy 241 CACAGAGAGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGGAATGCGAGAGACCTG 300
Db 318 CACAGAGAGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGGAATGCGAGAGACCTG 377
Qy 301 CAGAGCGTGTCCATCACTGTCTCAACGTCACCTCTGAACGACTCTGGCCCTCTACACCTGC 360
Db 378 CAGAGCGTGTCCATCACTGTCTCAACGTCACCTCTGAACGACTCTGGCCCTCTACACCTGC 437
Qy 361 AATGTGTCGGGAGTTGAGTTTGGCGCATCGGCCCTTTGTGAAGAGACCGCGCTG 420
Db 438 AATGTGTCGGGAGTTGAGTTTGGCGCATCGGCCCTTTGTGAAGAGACCGCGCTG 497
Qy 421 ATCCCTTTCAGAGTCCAGAGAGGCTGGAGAGACTTCACTCTGTGTCTCAGAAATC 480
Db 498 ATACCTTTCAGAGTCCAGAGAGGCTGGAGAGACTTCACTCTGTGTCTCAGAAATC 557
Qy 481 ATGATGTACATCTTCTGCTTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 558 ATGATGTACATCTTCTGCTTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
Qy 541 TACAGAAAGTCTCAAAAGCGGAGAGAGCGACCCCAAGAAACGCGTCTGACTACCTTGC 600
Db 618 TACAGAAAGTCTCTAAGCGCGAAGAGAGCGACCCCAAGAAATGCGTCTGACTACCTTGC 677
Qy 601 ATCCCTTTCAGAGTCCAGAGAGACTCTGCGGTACCGAGTGGAGGAATA 647
Db 678 ATCCCTTTCAGAGTCCAGAGAGACTCTGCGGTACCGAGTGGAGGAATA 724
```

## RESULT 10

AF378093

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITILE

JOURNAL

MEDLINE

PUBMED

AUTHORS

JOURNAL

FEATURES

source

AF378093 3107 bp mRNA linear ROD 11-SEP-2002  
Rattus norvegicus sodium channel beta 3 subunit (SCN3B) mRNA,  
complete cds.  
AF378093  
AF378093.1 GI:14165175  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 3107)  
Cui, Y., Curtis, R., Lawson, D., Gilbride, K., Ge, P., Distefano P.S.,  
Silos-Santiago, I., Catterall W.A. and Scheuer, T.  
Differential modulation of sodium channel gating and persistent  
sodium currents by the beta1, beta2, and beta3 subunits  
Mol. Cell. Neurosci. 18 (5), 570-580 (2001)  
21919106  
11922146  
2 (bases 1 to 3107)  
Curtis, R.  
Direct Submission  
Submitted (05-MAY-2001) Neurobiology, Millennium Pharmaceuticals  
Inc., 75 Sidney Street, Cambridge, MA 02139, USA  
Location/Qualifiers  
1. .3107  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/tissue\_type="dorsal root ganglion"

gene /dev stage="adult"  
1. .3107  
/gene="SCN3B"  
77. .724  
CDS /gene="SCN3B"  
/codon\_start=1  
/product="sodium channel beta 3 subunit"  
/protein\_id="AAK55415.1"  
/db\_xref="GI:14165176"  
/translation="MPAFNRLPLASLVLIYVVRVCPVCPVEPSEAVQGNPKLR  
CISMKKEEVEATVYVWFGKGLFLIYVRNGHVEFPGRLQWNGSKQLQD  
VSIIVLTLNDSGLYCNVSRPEFRRPFTVTRLIPLRVTEAGEDFTSVSEI  
MNYILLVFLMLFIEMLYCYRKYKAEAEQAENASDYLALPSENKENSVPVEE"

BASE COUNT 830 a 748 c 726 g 803 t  
ORIGIN  
Query Match 82.8%; Score 536.6; DB 10; Length 3107;  
Best Local Similarity 89.3%; Pred. No. 4.9e-116;  
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
Qy 1 ATGCTGCTCAATAGATTGTTCCCTGCTCTCTCGTCTATCTACTGGTCACT 60  
Db 77 ATGCTGCTCAATAGATTGTTCCCTGCTCTCTCGTCTATCTACTGGTCACT 136  
Qy 61 GTCTGCTTCCCTGCTGTTGGAAGTGCCTCGAGACGGAGCCGTCGAGGCAACCC 120  
Db 137 GTCTGCTTCCCTGCTGTTGGAAGTGCCTCGAGACGGAGCCGTCGAGGCAACCC 196  
Qy 121 ATGAGCTGCTGCTATCTCTGATGAGAGAGAGAGGAGTGGAGGCCACCACTGGT 180  
Db 197 ATGAGCTGCTGCTATCTCTGATGAGAGAGAGAGGAGTGGAGGCCACCACTGGT 256  
Qy 181 GAATGGTCTACAGCCGAGGCGGTAAAGTTCCTTATTTACGATATCGAATGGC 240  
Db 257 GAGTGGTCTACAGCCGAGGCGGTAAAGTTCCTTATTTACGATATCGAATGGC 316  
Qy 241 CACAGAGGTGGAGAGCCCTTTCCAGGGCGCTCGAGTGGAGTGGAGCAAGACCTG 300  
Db 317 CACAGAGGTGGAGAGCCCTTTCCAGGGCGCTCGAGTGGAGTGGAGCAAGACCTG 376  
Qy 301 CAGGACGTGCTCATCTGCTGCTCAACCTGCTGTAACAGCTCTGGCTCTACACCTGC 360  
Db 377 CAGGACGTGCTCATCTGCTGCTCAACCTGCTGTAACAGCTCTGGCTCTACACCTGC 436  
Qy 361 AATGTGCTCCGAGGTTTGAATTTGAGGCGCATCGGCCCTTTGTGAAGACACGCGCTG 420  
Db 437 AATGTGCTCCGAGGTTTGAATTTGAGGCGCATCGGCCCTTTGTGAAGACACGCGCTG 496  
Qy 421 ATCCCTTAAGAGTCAACGAGGAGCTCGAGAGGACTTCACTCTGTGGTCTCAGAAATC 480  
Db 497 ATACCTTTGCGAGTCACTGAAGGCGGAGAGACTTCACTCTGTGGTCTCAGAAATC 556  
Qy 481 ATGATGATACATCTTCTGGTCTTCTGCTACCTGCTGGCTGCTCATCGAGATATATGG 540  
Db 557 ATGATGATACATCTTCTGGTCTTCTGCTACCTGCTGGCTGCTCATCGAGATATATGG 616  
Qy 541 TACAGAAAGTCTCAAAAGCCGAGAGGAGGAGCCCAAGAAACCGCTCTGACTACCTGCC 600  
Db 617 TACAGAAAGTCTCAAAAGCCGAGAGGAGGAGCCCAAGAAACCGCTCTGACTACCTGCC 676  
Qy 601 ATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGGAGGAATA 647  
Db 677 ATCCCTTACAGAAACAGAGAACTCTGCGGTACCACTGGAGGAATA 723

RESULT 11  
AX047984  
LOCUS AX047984 3108 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 1 from Patent WO0069912.  
ACCESSION AX047984  
VERSION AX047984.1 GI:11876881  
KEYWORDS Rattus sp.

ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1  
AUTHORS Curtis R.A.  
TITLE Gene encoding a sodium channel beta-3 subunit protein  
JOURNAL Patent: WO 0069912-A 1 23-NOV-2000;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
source  
1. .3108  
/organism="Rattus sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10118"  
BASE COUNT 830 a 748 c 727 g 803 t  
ORIGIN  
Query Match 82.8%; Score 536.6; DB 6; Length 3108;  
Best Local Similarity 89.3%; Pred. No. 4.9e-116;  
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
Qy 1 ATGCTGCTCAATAGATTGTTCCCTGCTCTCTCGTCTATCTACTGGTCACT 60  
Db 78 ATGCTGCTCAATAGATTGTTCCCTGCTCTCTCGTCTATCTACTGGTCACT 137  
Qy 61 GTCTGCTTCCCTGCTGTTGGAAGTGCCTCGAGACGGAGCCGTCGAGGCAACCC 120  
Db 138 GTCTGCTTCCCTGCTGTTGGAAGTGCCTCGAGACGGAGCCGTCGAGGCAATCC 197  
Qy 121 ATGAGCTGCTGCTATCTCTGATGAGAGAGAGAGTGGAGGCCACCACTGGT 180  
Db 198 ATGAGCTGCTGCTATCTCTGATGAGAGAGAGAGTGGAGGCCACCACTGGT 257  
Qy 181 GAATGGTCTACAGCCGAGGCGGTAAAGTTCCTTATTTACGATATCGAATGGC 240  
Db 258 GAGTGGTCTACAGCCGAGGCGGTAAAGTTCCTTATTTACGATATCGAATGGC 317  
Qy 241 CACAGAGGTGGAGAGCCCTTTCCAGGGCGCTCGAGTGGAGTGGAGCAAGACCTG 300  
Db 318 CACAGAGGTGGAGAGCCCTTTCCAGGGCGCTCGAGTGGAGTGGAGCAAGACCTG 377  
Qy 301 CAGGACGTGCTCATCTGCTGCTCAACCTGCTGTAACAGCTCTGGCTCTACACCTGC 360  
Db 378 CAGGACGTGCTCATCTGCTGCTCAACCTGCTGTAACAGCTCTGGCTCTACACCTGC 437  
Qy 361 AATGTGCTCCGAGGTTTGAATTTGAGGCGCATCGGCCCTTTGTGAAGACACGCGCTG 420  
Db 438 AATGTGCTCCGAGGTTTGAATTTGAGGCGCATCGGCCCTTTGTGAAGACACGCGCTG 497  
Qy 421 ATCCCTTAAGAGTCAACGAGGAGCTCGAGAGGACTTCACTCTGTGGTCTCAGAAATC 480  
Db 498 ATACCTTTGCGAGTCACTGAAGGCGGAGAGACTTCACTCTGTGGTCTCAGAAATC 557  
Qy 481 ATGATGATACATCTTCTGGTCTTCTGCTACCTGCTGGCTGCTCATCGAGATATATGG 540  
Db 558 ATGATGATACATCTTCTGGTCTTCTGCTACCTGCTGGCTGCTCATCGAGATATATGG 617  
Qy 541 TACAGAAAGTCTCAAAAGCCGAGAGGAGGAGCCCAAGAAACCGCTCTGACTACCTGCC 600  
Db 618 TACAGAAAGTCTCAAAAGCCGAGAGGAGGAGCCCAAGAAACCGCTCTGACTACCTGCC 677  
Qy 601 ATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGGAGGAATA 647  
Db 678 ATCCCTTACAGAAACAGAGAACTCTGCGGTACCACTGGAGGAATA 724

RESULT 12  
AX048004  
LOCUS AX048004 645 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 21 from Patent WO0069912.  
ACCESSION AX048004  
VERSION AX048004.1 GI:11876882  
KEYWORDS

SOURCE  
ORGANISM Lepus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Lagomorpha; Leporidae; Lepus.  
REFERENCE  
AUTHORS Curtis, R.A.  
TITLE Gene encoding a sodium channel beta-3 subunit protein  
JOURNAL Patent: WO 0069912-A 21 23-NOV-2000;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..645  
/organism="Lepus sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:137772"  
BASE COUNT 155 a 154 c 181 g 155 t  
ORIGIN  
Query Match 82.5%; Score 534.6; DB 6; Length 645;  
Best Local Similarity 89.3%; Pred. No. 1.6e-115;  
Matches 576; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
Qy 1 ATGCTGCTTCAATAGATTGTTCCCTGCTTCTCGTGCTTATCTACTGGTCACT 60  
Db |||||||  
1 ATGCTGCTTCAACAGATTGTTCCCTGCTTCTAGTGCTCACTACTGGTCACTA 60  
Qy 61 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACGAGGCGGTGCGAGCAACCCC 120  
Db |||||||  
61 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACGAGGCGGTGCGAGCAATCCC 120  
Qy 121 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Db |||||||  
121 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Qy 181 GAATGTTCTACAGCCGAGGCGGTAAAGATTTCCTTATTAGAGTATCGGAATGGC 240  
Db |||||||  
181 GAGTGGTCTACAGCCCTGAGGCGGTAAAGATTTCCTTATTAGAGTATCGGAATGGC 240  
Qy 241 CACGAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
Db |||||||  
241 CACGAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
Qy 301 CAGGAGGTGCTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db |||||||  
301 CAGGAGGTGCTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Qy 361 AATGTGTCGGGAGTTTGTGTTGAGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 420  
Db |||||||  
361 AATGTGTCGGGAGTTTGTGTTGAGCGCATCGGCCCTTTGTGAAGACGACGACTG 420  
Qy 421 ATCCCTTAAGAGTACCGAGGCTGAGAGGACTTCACCTCTGTGGTCTCAGAAATC 480  
Db |||||||  
421 ATACCTTTGCGAGTCACTGAAGAGCGCGGGAAGACTTCACCTCGGTGCTCGGAATC 480  
Qy 481 ATGATGTACATCTTCTGCTTCTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db |||||||  
481 ATGATGTACATCTTCTGCTTCTCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Qy 541 TACGAAAGGTCTCAAAAGCCGAGAGCAGCCCAAGAAACCGGCTCTGACTCTGCC 600  
Db |||||||  
541 TACGAAAGGTCTCTAAGCCGAGAGGAGCAGCAGGAAATGCTCTGACTACCTTGGT 600  
Qy 601 ATCCATCTGAGACAGGAGACTCTCGGTACGAGTGGAGAA 645  
Db |||||||  
601 ATCCCTTCAGAGAACAGGAGAACTCTGTGTACTGTGGAGGAA 645

RESULT 13  
BD059018 BD059018 471 bp DNA linear PAT 27-AUG-2002  
LOCUS  
DEFINITION Secreted expressed sequence tags (sESTs).  
ACCESSION BD059018  
VERSION BD059018.1 GI:22604624  
KEYWORDS JP 2001519666-A/873.

SOURCE  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS Jacoby, K., Mccoy, J.M., Lavallie, E.R., Racine, L.A., Merberg, D.,  
Treacy, M., Spaulding, V. and Agostino, M.J.  
TITLE Secreted expressed sequence tags (sESTs)  
JOURNAL Patent: JP 2001519666-A 873 23-OCT-2001;  
GENETICS INSTITUTE INC  
COMMENT PN JP 2001519666-A/873  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACINE, PI  
DAVID MERBERG,  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
FEATURES  
source  
1..471  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
BASE COUNT 93 a 126 c 143 g 108 t 1 others  
ORIGIN  
Query Match 65.1%; Score 422; DB 6; Length 471;  
Best Local Similarity 99.5%; Pred. No. 5.8e-89;  
Matches 433; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
Qy 1 ATGCTGCTTCAATAGATTGTTCCCTGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCT 60  
Db |||||||  
29 ATGCTGCTTCAATAGATTGTTCCCTGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCT 88  
Qy 61 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACGAGGCGGTGCGAGGCAACCCC 120  
Db |||||||  
89 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACGGA-GCCGTGCGAGGCAACCCC 147  
Qy 121 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Db |||||||  
148 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 207  
Qy 181 GAATGTTCTACAGGCGGCGGTAAAGATTTCCTTATTACGAGTATCGGAATGGC 240  
Db |||||||  
208 GAATGTTCTACAGGCGGCGGTAAAGATTTCCTTATTACGAGTATCGGAATGGC 267  
Qy 241 CACGAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
Db |||||||  
268 CACGAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327  
Qy 301 CAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db |||||||  
328 CAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387  
Qy 361 AATGTGTCGGGAGTTTGTGTTGAGGCGCATCGGCCCTTTGTGAAGAGAGCGCGCTG 420  
Db |||||||  
388 AATGTGTCGGGAGTTTGTGTTGAGGCGCATCGGCCCTTTGTGAAGAGAGCGCGCTG 447  
Qy 421 ATCCCTTAAGAGTC 435  
Db |||||||  
448 ATCCCTTAAGAGTC 462

RESULT 14  
AP000682/c AP000682/c 127347 bp DNA linear HTG 16-AUG-2000  
LOCUS  
DEFINITION Homo sapiens chromosome 11 clone CMB9-32A1 map 11q24, WORKING DRAFT  
SEQUENCE, 8 unordered pieces.  
ACCESSION AP000682

AP000682.3 GI:9844967  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,  
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Homo sapiens 127,347 genomic DNA of 11q24  
Published Only in Database (1999)  
2 (bases 1 to 127347)  
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,  
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (08-NOV-1999) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan  
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,  
Tel: 81-42-778-9923, Fax: 81-42-778-9924)  
On Aug 18, 2000 this sequence version replaced gi:8119870.  
----- Genome Center  
Center: RIKEN Genomic Sciences Center (GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: HumDraft11  
Center clone name: CMB9-32A1  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap; version 0.950329  
Consensus quality: 124981 bases at least Q40  
Consensus quality: 125815 bases at least Q30  
Consensus quality: 126286 bases at least Q20  
Insert size: 126647; sum-of-contigs  
Quality coverage: 12.71x in Q20 bases; sum-of-contigs  
-----  
NOTE: This is a 'working draft' sequence. It currently consists of  
8 contigs. The true order of the pieces is not known and their  
order in this sequence record is arbitrary. Gaps between the  
contigs are represented as runs N, but the exact sizes of the gaps  
are unknown. This record will be updated with the finished sequence  
as soon as it is available and the accession number will be  
preserved  
1 33532 contig of 33532 bp in length  
33633 59223 contig of 25591 bp in length  
59324 83007 contig of 23684 bp in length  
83108 100196 contig of 17089 bp in length  
100297 112313 contig of 12017 bp in length  
112414 121043 contig of 8630 bp in length  
121144 125917 contig of 1330 bp in length  
126018 127347 contig of 1330 bp in length  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 33532 contig of 33532 bp in length  
33633 33632 gap of 100 bp  
33633 59223 contig of 25591 bp in length  
59224 59323 gap of 100 bp  
59324 83007 contig of 23684 bp in length  
83008 83107 gap of 100 bp  
83108 100196 contig of 17089 bp in length  
100197 100296 gap of 100 bp  
100297 112313 contig of 12017 bp in length  
112314 112414 gap of 100 bp  
112414 121043 contig of 8630 bp in length

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE  
JOURNAL

REFERENCE  
AUTHORS

TITLE  
JOURNAL

COMMENT

112143: gap of 100 bp  
121044 121143: gap of 100 bp  
125917: contig of 4774 bp in length  
126017: gap of 100 bp  
126018 127347: contig of 1330 bp in length.

FEATURES  
source  
1..127347  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q24"  
/clone="CMB9-32A1"  
1..33532  
/note="assembly\_fragment"  
33633..59223  
/note="assembly\_fragment"  
59324..83007  
/note="assembly\_fragment"  
83108..100196  
/note="assembly\_fragment clone\_end:SP6 vector\_side:left"  
100297..112313  
/note="assembly\_fragment"  
112414..121043  
/note="assembly\_fragment clone\_end:T7 vector\_side:right"  
121144..125917  
/note="assembly\_fragment"  
126018..127347  
/note="assembly\_fragment"  
BASE COUNT 36325 a 27508 c 27825 g 34989 t 700 others  
ORIGIN

Query Match 35.4%; Score 229.6; DB 2; Length 127347;  
Best Local Similarity 96.3%; Pred. No. 1.1e-43;  
Matches 235; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 212 ATTTCCTTATTTACGATATCGAATGCCACAGGAGGTGGAGAGCCCTTTTCAGGGGC 271  
|||  
DB 17650 ATCTGCAGATTACGATATCGAATGCCACAGGAGGTGGAGAGCCCTTTTCAGGGGC 17591  
QY 272 GCCTGCAGTGAATGGCGAGCAAGACCTGCGAGGAGTGTCCATCTGCTCAACGTCA 331  
|||  
DB 17590 GCCTGCAGTGAATGGCGAGCAAGACCTGCGAGGAGTGTCCATCTGCTCAACGTCA 17531  
QY 332 CTCTGAACGACTCTGGCCCTCTACACCTGCAATGTCTCCGGAGTTTCAGTTTCAGGGGC 391  
|||  
DB 17530 CTCTGAACGACTCTGGCCCTCTACACCTGCAATGTCTCCGGAGTTTCAGTTTCAGGGGC 17471  
QY 392 ATCGGCCCTTTGTGAAGACGACGCGGCTGTATCCCTTAAGAGTCAACGAGAGGCTGAG 451  
|||  
DB 17470 ATCGGCCCTTTGTGAAGACGACGCGGCTGTATCCCTTAAGAGTCAACGAGAGGCTGAG 17411  
QY 452 AGGA 455  
|||  
DB 17410 CTGA 17407

RESULT 15  
AC024604/c  
LOCUS Homo sapiens chromosome Chromosome 10 clone RF11-331G19, WORKING  
DEFINITION DRAFT SEQUENCE, 22 unordered pieces.  
ACCESSION AC024604  
VERSION AC024604.3 GI:8389428  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_CANCELLED.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 172546)  
TITLE Smith, D.R.  
JOURNAL Genome Therapeutics Corporation Sequencing Center: Human Genome  
Sequence Data  
Unpublished

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

2 (bases 1 to 172546)  
Smith, D.R.  
Direct Submission  
Submitted (01-MAR-2000) Genome Therapeutics Corporation, 100 Beaver  
Street, Waltham, MA 02453, USA  
On Jun 9, 2000 this sequence version replaced gi:7549605.  
Genome Center  
Center: Genome Therapeutics Corporation  
Center code: GTC  
Web site: <http://www.genomecorp.com/>  
Contact: gtc-seqcenter@genomecorp.com  
----- Project Information -----  
Center project name: Hg202  
----- Summary Statistics -----  
Sequencing vector: N/A  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 990315  
Consensus quality: 152033 bases at least Q40  
Consensus quality: 163535 bases at least Q30  
Consensus quality: 165256 bases at least Q20  
Insert size: 170446; sum-of-contigs  
Quality coverage: 4.3x in Q20 bases; sum-of-contigs

FEATURES \* 135419 172546: contig of 37128 bp in length.  
Location/Qualifiers  
1. 172546  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="Chromosome 10"  
/clone="RP11-331G19"  
/clone\_lib="RPC1-11"  
BASS COUNT 47530 a 37112 c 38612 g 46556 t 2136 others  
ORIGIN  
Query Match 35.4%; Score 229.6; DB 2; Length 172546;  
Best Local Similarity 96.3%; Pred. No. 1.1e-43;  
Matches 235; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 212 ATTCCTTATTTACGAGTATCGGATGGCCACGAGAGGTGGAGAGCCCTTTACGGGC 271  
DB 54796 ATCTGCAGATTTACGAGTATCGGATGGCCACGAGAGGTGGAGAGCCCTTTACGGGC 54737  
QY 272 GCTGCGAGTGGGAATGGCAGCAAGGACCTGCAGGACGTGTCCATCACTGTGCTCAACGTCA 331  
DB 54736 GCTGCGAGTGGGAATGGCAGCAAGGACCTGCAGGACGTGTCCATCACTGTGCTCAACGTCA 54677  
QY 332 CTCTGAACGACACTCTGGCCCTCTACACTGCAATGTGTCCCGGAGTTTGAGTTTGGAGCGC 391  
DB 54676 CTCTGAACGACACTCTGGCCCTCTACACTGCAATGTGTCCCGGAGTTTGAGTTTGGAGCGC 54617  
QY 392 ATCGGCCCTTTGTGAAGACGACGCGCTGATCCCCCTTAAGAGTCAACGAGGAGGCTGGAG 451  
DB 54616 ATCGGCCCTTTGTGAAGACGACGCGCTGATCCCCCTTAAGAGTCAACGAGGAGGCTGGAG 54557  
QY 452 AGGA 455  
DB 54556 CTGA 54553  
Search completed: December 5, 2003, 20:50:55  
Job time : 2707.8 secs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 22 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
1339: contig of 1339 bp in length  
1439: gap of unknown length  
2789: contig of 1350 bp in length  
2889: gap of unknown length  
4562: contig of 1673 bp in length  
4662: gap of unknown length  
6344: contig of 1682 bp in length  
6444: gap of unknown length  
9221: contig of 2777 bp in length  
9321: gap of unknown length  
11997: contig of 2676 bp in length  
12097: gap of unknown length  
14867: contig of 2770 bp in length  
14967: gap of unknown length  
19919: contig of 4952 bp in length  
20019: gap of unknown length  
24125: contig of 4106 bp in length  
24225: gap of unknown length  
29444: contig of 5219 bp in length  
29544: gap of unknown length  
34568: contig of 5024 bp in length  
34668: gap of unknown length  
39414: contig of 4746 bp in length  
39514: gap of unknown length  
45885: contig of 6371 bp in length  
45985: gap of unknown length  
50997: contig of 5012 bp in length  
51097: gap of unknown length  
59143: contig of 8046 bp in length  
59243: gap of unknown length  
67779: contig of 8536 bp in length  
67879: gap of unknown length  
76429: contig of 8550 bp in length  
76529: gap of unknown length  
86704: contig of 10175 bp in length  
86804: gap of unknown length  
103147: contig of 16343 bp in length  
103247: gap of unknown length  
116343: contig of 13096 bp in length  
116443: gap of unknown length  
135318: contig of 18875 bp in length  
135418: gap of unknown length  
135319

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:12:31 ; Search time 244.092 Seconds  
(without alignments)  
7166.298 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_376\_1023

Perfect score: 648

Sequence: 1 atgcttgccttcattagatt.....cggtagcagtgagggaatag 648

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	648	100.0	1261	21	Human beta3 cDNA.
2	646.4	99.8	1510	22	Human novel sodium
3	646.4	99.8	4052	24	Human signal trans
4	585	90.3	1045	22	Human polynucleoti
5	570.2	88.0	978	22	Human EST-derived
6	536.6	82.8	2220	21	Rat beta3 subunit
7	536.6	82.8	2632	22	Rat sodium channel
8	536.6	82.8	3108	22	Rat sodium channel

9	534.8	82.5	1195	23	AAS86764	DNA encoding novel
10	534.6	82.5	645	22	AAC90601	Rat sodium channel
11	422	65.1	471	20	AAV86895	EST clone BMA. Ho
12	392.6	60.6	3531	23	AAS86763	DNA encoding novel
13	224.4	34.6	621	23	AAS86762	DNA encoding novel
14	148.2	22.9	657	22	AAC90603	Rabbit sodium chan
15	142.2	21.9	1414	25	ABZ23837	Human voltage-gate
16	141	21.8	407	22	ABA08942	Human voltage-gate
17	138.6	21.4	407	22	AAK53329	Human polynucleoti
18	138.6	21.4	407	22	AAK53329	Rat sequence diffe
19	81.4	12.6	974	24	ABK63747	Human beta3A sodiu
20	73	11.3	850	22	AAK53329	Rat sodium channel
21	73	11.3	850	22	AAK53329	Sodium channel bet
22	47	7.3	509	22	AAK53329	Human polynucleoti
23	45.4	7.0	1645	21	AAK53329	Maize geranylgeran
24	42.2	6.5	500	21	AAK53329	Zea mays DNA fragm
25	39.2	6.0	5973	24	ABK63747	DNA encoding human
26	39.2	6.0	6202	25	ABK63747	Novel human cDNA s
27	39.2	6.0	7584	25	ABK63747	Human secretory po
28	38.8	6.0	1299	23	ABL10322	Drosophila melanog
29	38.8	6.0	3787	23	ABL10322	Drosophila melanog
30	38.6	6.0	1179	23	AAK54061	Pseudomonas aerugi
31	38.6	6.0	2423	23	ABL24082	Drosophila melanog
32	38.6	6.0	4166	23	ABL24084	Drosophila melanog
33	38	5.9	1122	21	AAK30433	Human ACAM#6 PCR p
34	37.8	5.8	1234	23	ABL28305	Drosophila melanog
35	37.4	5.8	6217	24	ABL32239	Human immune syste
36	37.2	5.7	721	22	AAH08615	Human cDNA clone (
37	37.2	5.7	1642	22	AAH16770	Human cDNA sequenc
38	37.2	5.7	2704	23	ABL13675	Drosophila melanog
39	37.2	5.7	12276	23	ABL13674	Drosophila melanog
40	37	5.7	1558	22	AAK76859	Human secreted pro
41	36.6	5.6	1548	22	AAK76859	Human ovarian and
42	36.6	5.6	1548	22	AAK76859	Human reproductive
43	36.6	5.6	1548	22	AAK76859	Human immune/haema
44	36.6	5.6	1548	22	AAK76859	Drosophila melanog
45	36.6	5.6	6712	23	ABL10229	Drosophila melanog

ALIGNMENTS

RESULT 1	
AAC67837	
ID AAC67837 standard; cDNA; 1261 BP.	
XX AAC67837;	
AC AAC67837;	
DT 15-FEB-2001 (first entry)	
XX Human beta3 cDNA.	
DE DE	
XX Human; beta sub-unit; beta3; analgesic; anticonvulsant;	
KW cerebroprotective; vasotropic; cardiac; norepinephrine;	
KW dermatological; gene therapy; voltage-gated sodium channel; pain;	
KW epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;	
KW familial nonchromaffin paraganglioma; phenylketonuria;	
KW Charcot Marie Tooth disease; ss.	
XX Homo sapiens.	
XX WO200063367-A1.	
PN WO200063367-A1.	
XX 26-OCT-2000.	
PD 26-OCT-2000.	
XX 24-FEB-2000; 2000WO-EP01783.	
XX 15-APR-1999; 99US-0129473.	
XX (WARN ) WARNER LAMBERT CO.	
PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.	
XX Cox P, Dixon A, Jackson A, Morgan K;	



Db 453 CACAGGAGGTGGAGAGCCCTTTTCAGGGGGCGCTGCTGAGTGAATGCGCAGAGGACCTG 512  
QY 301 CAGGAGCTTCATCATCTGCTCAACGTCACCTCTCAACGACTCTGGCCCTCTACACCTGC 360  
Db 513 CAGGAGGTGTCCATCATCTGTGCTCAACGTCACCTCTCAACGACTCTGGCCCTCTACACCTGC 572  
QY 361 AATGTCTCCGGGAGTTTGAAGTTTGAAGGCGCATCGGCCCTTTTGTGAAGACGACGCGGCTG 420  
Db 573 AATGTCTCCGGGAGTTTGAAGTTTGAAGGCGCATCGGCCCTTTTGTGAAGACGACGCGGCTG 632  
QY 421 ATCCCTCTAAGAGTCACCGAGGAGGCTGGAGAGGACTTCACTCTCTGTGTCTCAGAAATC 480  
Db 633 ATCCCTCTAAGAGTCACCGAGGAGGCTGGAGAGGACTTCACTCTCTGTGTCTCAGAAATC 692  
QY 481 ATGATGTACATCTCTGCTGCTTCTCCTCACCCTGCTGCTGCTCATCGAGATGATATATTC 540  
Db 693 ATGATGTACATCTCTGCTGCTTCTCCTCACCCTGCTGCTGCTCATCGAGATGATATATTC 752  
QY 541 TACAGAAAGTCTCAAAAGCCGAGAGGCGCCCAAGAAACCGCTCTGACTACCTTGCC 600  
Db 753 TACAGAAAGTCTCAAAAGCCGAGAGGCGCCCAAGAAACCGCTCTGACTACCTTGCC 812  
QY 601 ATCCATCTGAGAACAGAGGAACTCTGCGGTACCGTACGAGGAATAG 648  
Db 813 ATCCATCTGAGAACAGAGGAACTCTGCGGTACCGTACGAGGAATAG 860

## RESULT 3

ABA93727 ID ABA93727 standard; cDNA; 4052 BP.

XX ABA93727;

XX 30-APR-2002 (first entry)

XX Human signal transduction cDNA clone amy2\_2f19.

XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;

XX gene therapy; ss.

XX Homo sapiens.

XX WO200198454-A2.

XX 27-DEC-2001.

XX 25-APR-2001; 2001WO-1B02050.

XX 25-APR-2000; 2000US-199380P.

XX (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX Wiemann S;

XX WPT; 2002-055860/07.

XX P-PSDE; ABB05689.

XX Human cDNA sequences and clones derived from human fetal brain, fetal  
XX kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic  
XX screening and therapy -

XX Claim 1; Page 174-175; 61pp; English.

XX The present invention describes assemblages and computer readable media  
XX comprising novel human cDNA sequences and clones derived from human  
XX foetal brain, foetal kidney, melanoma, testis and amygdala cDNA  
XX libraries. ABA93702 to ABA93766 represent human cDNA sequences from the  
XX present invention which encode the proteins given in ABB05662 to  
XX ABB05729. The human cDNA sequences and clones can be used in gene  
XX therapy. The clones may be used in a variety of applications, for  
XX example they may be used in profiling assays, for providing large arrays  
XX of human genetic material for implementing large-scale screening  
XX strategies and for treating diseases via gene therapy procedures.

XX SQ Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 other;

Query Match 99.8%; Score 646.4; DB 24; Length 4052;

Best Local Similarity 99.8%; Pred. No. 1e-166; Mismatches 1; Indels 0; Gaps 0;

Matches 647; Conservative 0;

QY 1 ATGCTGCTTCAATAGATTTGTTCCCTGGCTTCTCTGCTTATCTACTTGGTCACT 60

Db 804 ATGCTGCTTCAATAGATTTGTTCCCTGGCTTCTCTGCTTATCTACTTGGTCACT 863

QY 61 GTCTGCTTCCCTGTGTGTGAAGTGCCTCGGAGACGAGGCGCTGAGGCAACCC 120

Db 864 GTCTGCTTCCCTGTGTGTGAAGTGCCTCGGAGACGAGGCGCTGAGGCAACCC 923

QY 121 ATGAGAGTCCGCTGCATCTCTGCTGATGAGAGAGGAGGTGAGGCCACCAACCGTGTG 180

Db 924 ATGAGAGTCCGCTGCATCTCTGCTGATGAGAGAGGAGGTGAGGCCACCAACCGTGTG 983

QY 181 GAATGCTTCTCAGAGCCCGAGGCGGTAAAGTTTCTTTATTTACGAGTATCGAATGTC 240

Db 984 GAATGCTTCTCAGAGCCCGAGGCGGTAAAGTTTCTTTATTTACGAGTATCGAATGTC 1043

QY 241 CACAGAGGTGAGAGGCGCTTTTCAGGGCGCTGAGTGAATGCGCAGGACCTG 300

Db 1044 CACAGAGGTGAGAGGCGCTTTTCAGGGCGCTGAGTGAATGCGCAGGACCTG 1103

QY 301 CAGAGGTGCTCATCTGCTCAACGTCACCTCTGAACGACTCTGGCTCTACACCTGC 360

Db 1104 CAGAGGTGCTCATCTGCTCAACGTCACCTCTGAACGACTCTGGCTCTACACCTGC 1163

QY 361 AATGTGTCGGGAGTTTGAAGTTCAGGCGCATTCGGCCCTTTGTGAAGACGACGCGGCTG 420

Db 1164 AATGTGTCGGGAGTTTGAAGTTCAGGCGCATTCGGCCCTTTGTGAAGACGACGCGGCTG 1223

QY 421 ATCCCTCTAAGAGTCAACGAGGCGCTGAGAGGACTTCACTCTGCTCTCAGAAATC 480

Db 1224 ATCCCTCTAAGAGTCAACGAGGCGCTGAGAGGACTTCACTCTGCTCTCAGAAATC 1283

QY 481 ATGATGTACATCTCTGCTGCTTCTCTCACTGCTGCTCATCGAGATGATATATTC 540

Db 1284 ATGATGTACATCTCTGCTGCTTCTCTCACTGCTGCTCATCGAGATGATATATTC 1343

QY 541 TACAGAAAGTCTCAAAAGCCGAGAGGCGAGCCCAAGAAACCGCTCTGACTACCTTGC 600

Db 1344 TACAGAAAGTCTCAAAAGCCGAGAGGCGAGCCCAAGAAACCGCTCTGACTACCTTGC 1403

QY 601 ATCCATCTGAGAACAGGAGGACTCTGCGGTACCGTACGAGGAATAG 648

Db 1404 ATCCATCTGAGAACAGGAGGACTCTGCGGTACCGTACGAGGAATAG 1451

## RESULT 4

AAK52345

ID AAK52345 standard; cDNA; 1045 BP.

XX AAK52345;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 890.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX





## RESULT 7

AAC90602  
ID AAC90602 standard; DNA; 2632 BP.  
XX  
AC AAC90602;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Rat sodium channel beta3 protein Alrxa94h5 related sequence.  
XX  
KW Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;  
KW neurodegenerative disorder; mood disorder; muscle contraction; ds.  
XX  
XX  
OS Rattus sp.  
XX  
PN WO200069912-A1.  
XX  
PD 23-NOV-2000.  
XX  
PF 12-MAY-2000; 2000WO-US13144.  
XX  
PR 14-MAY-1999; 99US-0134198.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Curtis RAJ;  
XX  
DR WPI; 2001-122743/13.  
XX  
PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal  
PT root ganglion cDNA library for use in chromosome mapping, forensic  
PT medicine, monitoring clinical trials and therapeutics -  
XX  
PS Claim 1; Page 41-42; 145pp; English.  
XX  
CC The present invention provides the protein and coding sequences of the  
CC rat sodium channel beta3 protein, designated Alrxa94h5. This protein is  
CC involved in the generation of pain and other sensory or perceptive nerve  
CC impulses, in the establishment and endurance of mood, neurodegenerative  
CC and sleep disorders, and in the control of muscle contraction, including  
CC movements such as the heartbeat, digestion and vascular tone. The  
CC sequences can be used in predictive medicine, screening and diagnostic  
CC assays, and in pharmacogenomics.  
XX  
SQ Sequence 2632 BP; 693 A; 636 C; 611 G; 692 T; 0 other;  
Query Match 82.8%; Score 536.6; DB 22; Length 2632;  
Best Local Similarity 89.3%; Pred. No. 1e-136;  
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 1 ATGCGCTGCTCAATAGATTGTTCCCTGGCTTCTCTGCTGCTTATCTACTGGTCACT 60  
DB 78 ATGCGCTGCTTCAAGATGCTTCCCTAGCTTCTAGTCTATCTACTGGTCACT 137  
QY 61 GTCTGCTTCCCTGTGTGTGAAGTCCCTCGGACGAGCGCGTGCAGGCGCAACCCC 120  
DB 138 GTCTGCTTCCCTGTGTGTGAAGTCCCTCGGACGAGCGCGTGCAGGCGCAATCC 197  
QY 121 ATGAAGCTGCTGCTATCTCTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 198 ATGAAGCTGAGTGCTATCTCTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 257  
QY 181 GAATGTTCTACAGGCGCGAGGCGGTAAAGATTCTTATTTACGAGTATCGAATGCG 240  
DB 258 GAGTGTCTTACAGGCGCTGAGGCGGTAAAGATTCTTATTTATGAGTATCGAATGCG 317  
QY 241 CACCAGAGGTGAG 300  
DB 318 CACCAG 377  
QY 301 CAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

DB 378 CAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437  
QY 361 AATGTGTCCTGGAGTTTGAAGTTCGAGCGCATCGGCCCTTTGTGAAGACGACGCGGTG 420  
DB 438 AATGTGTCCTGGAGTTTGAAGTTCGAGCGCATCGGCCCTTTGTGAAGACGACGAGACTG 497  
QY 421 ATCCCTTAAGAGTCAAG 480  
DB 498 ATACCTTTGCGAGTCACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 557  
QY 481 ATGATGATACATCTTCTGCTTCTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB 558 ATGATGATACATCTTCTGCTTCTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617  
QY 541 TACAGAAAGTCTCAAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
DB 618 TACAGAAAGTCTCAAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 677  
QY 601 ATCCCTTCAAG 647  
DB 678 ATCCCTTCAAG 724

## RESULT 8

AAC90600  
ID AAC90600 standard; cDNA; 3108 BP.

XX

AC AAC90600;

XX 13-MAR-2001 (first entry)

XX DE Rat sodium channel beta3 protein Alrxa94h5 coding sequence.

XX KW Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;  
KW neurodegenerative disorder; mood disorder; muscle contraction; ss.

XX OS Rattus sp.

XX PN WO200069912-A1.

XX PD 23-NOV-2000.

XX PF 12-MAY-2000; 2000WO-US13144.

XX PR 14-MAY-1999; 99US-0134198.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Curtis RAJ;

XX DR WPI; 2001-122743/13.

XX DR P-PSDB; AAB50243, AAB50245.

XX PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal  
PT root ganglion cDNA library for use in chromosome mapping, forensic  
PT medicine, monitoring clinical trials and therapeutics -

XX PS Claim 1; Fig 1; 145pp; English.

XX CC The present invention provides the protein and coding sequences of the  
XX rat sodium channel beta3 protein, designated Alrxa94h5. This protein is  
XX involved in the generation of pain and other sensory or perceptive nerve  
XX impulses, in the establishment and endurance of mood, neurodegenerative  
XX and sleep disorders, and in the control of muscle contraction, including  
XX movements such as the heartbeat, digestion and vascular tone. The  
XX sequences can be used in predictive medicine, screening and diagnostic  
XX assays, and in pharmacogenomics.

XX SQ Sequence 3108 BP; 830 A; 748 C; 727 G; 803 T; 0 other;

Query Match

Best Local Similarity 82.8%; Score 536.6; DB 22; Length 3108;

Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

```
QY 1 ATGCGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCGTGTCTTATCTACTGCTCAGT 60
DB |||||
DB 78 ATGCGCTGCTTCAACAGATTGTTCCCTAGCTTCTAGTGTCTATCTACTGCTCAGA 137
QY 61 GTCTGCTTCCCTGTGTGTGTGAAGTGCCTCGAGAGCGGCGGTGAGGGCAACCC 120
DB |||||
DB 138 GTCTGCTTCCCTGTGTGTGAAGTGCCTCGAGAGCGGCGGTGAGGGCAATCCC 197
QY 121 ATGAAGCTGCGCTGCATCTCTCGCATGAAGAGAGAGGAGGTGGAGCCACACGCTGTG 180
DB |||||
DB 198 ATGAAGCTGAGTGTGATCTCTCTGATGAAGAGAGAGGAGGTGGAGCCACACACTGTGTG 257
QY 181 GAATGGTTCTACAGCCCGGAGCGGTTAAAGATTTCTTATTTACGAGTATCGGAATGCC 240
DB |||||
DB 258 GAGTGGTTCTACAGCCCTGAGGCGGTAAAGATTTCTTATATATAGTATCGGAATGCC 317
QY 241 CACCAAGAGGTGGAGAGCCCTTTCAGGGGCGCTCGAGTGAATGGAGCAGACCTG 300
DB |||||
DB 318 CACCAAGAGGTGGAGAGCCCTTTCAGAGCGGCTCTGCAAGTGAATGGAGCAAGACCTG 377
QY 301 CAGGACGTGTCCATCACTGTGTCTCAACGTCACTCTGAACGACTCTGGCTCTACACCTGC 360
DB |||||
DB 378 CAGGACGTATCCATCACTGTGTCTCAATGTCACTTTGATGACTCTGGCTCTACACAGC 437
QY 361 AATGTGTCCCGGAGTTGAGTTGAGCGCATCGGCCCTTTGTGAAGACACGCGCTG 420
DB |||||
DB 438 AATGTGTCCCGGAGTTGAGTTGAGCGCATCGGCCCTTTGTGAAGACACGAGACTG 497
QY 421 ATCCCTTCAAGAGTCAACCGAGGCTGGAGGAGCTTCACTCTGTGTGTCTCAGAAATC 480
DB |||||
DB 498 ATACCTTTGCGAGTCACTGAAGAGCGGAGAGACTTCACTCCCTGCTCTCGGAATC 557
QY 481 ATGATGTATCATCTTCTGTGTCTTCTCACCCTGTGGCTGCTCATCGAGATGATATATTC 540
DB |||||
DB 558 ATGATGTATCATCTTCTGTGTCTTCTCACCCTGTGGCTGCTTATGTGATGATCTATTCG 617
QY 541 TACAGAAAGGTCTCAAGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB |||||
DB 618 TACAGAAAGGTCTTCAAGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677
QY 601 ATCCCTTCAAGAGTCAACCGAGGAGCTTCACTCTGTGTGTGTCTCAGAAATC 647
DB |||||
DB 678 ATCCCTTCAAGAGTCAACCGAGGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 724

RESULT 9
AA886764
ID AA886764 standard; cDNA; 1195 BP.
AC AA886764;
XX
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #22568.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HvSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
```

```
XX WPI; 2001-639362/73.
DR P-PSDB; ABG22577.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID No 22568; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridization probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AA864197-BA894564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1195 BP; 253 A; 356 C; 381 G; 205 T; 0 other;
```

```
Query Match 82.5%; Score 534.8; DB 23; Length 1195;
Best Local Similarity 95.7%; Pred. No. 2.3e-136;
Matches 572; Conservative 0; Mismatches 22; Indels 4; Gaps 2;

QY 1 ATGCTCCCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGTCTTATCTACTGCTCAGT 60
DB 463 ATGCTCCCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGTCTTATCTACTGCTCAGT 522
QY 61 GTCTGCTTCCCTGTGTGTGTGAA--GTGCCCTCGAGAGCGGAGCGGCTGCA--GGGCAA 116
DB 523 GTCTGCTTCCCTGTGTGTGTGAAAGTCCCTTAGAAAAGGGGGCGGACGGGTAA 582
QY 117 CCCATGAAGCTGCGCTGTGATCTCTCTCATGAAGAGAGAGAGTGAAGGCCACCGGT 176
DB 583 CCCATGAAGCTGCGCTGTGATCTCTCTCATGAAGAGAGAGTGAAGGCCACCGGT 642
QY 177 GGTGGAATGTTCTTACAGGCCCGAGGGCGGTAAAGATTTCCTTATTTACGATATCGGAA 236
DB 643 GGTGGAATGTTCTTACAGGCCCGAGGGCGGTAAAGATTTCCTTATTTACGATATCGGAA 702
QY 237 TGGCCACAGAGAGTGGAGAGCGCCCTTTTCAGGGGCGCTTCAGTGAATGCGACAAAGGA 296
DB 703 TGGCCACAGAGTGGAGAGCGCCCTTTTCAGGGGCGCTTCAGTGAATGCGACAAAGGA 762
QY 297 CCTGCAGAGCTGTCCATCTGTCTCAACGTCATCTCTGAACGACTCTGCGCTCTACAC 356
DB 763 CCTGCAGAGCTGTCCATCTGTCTCAACGTCATCTCTGAACGACTCTGCGCTCTACAC 822
QY 357 CTGCAATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGAGCAGCGC 416
DB 823 CTGCAATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGAGCAGCGC 882
QY 417 GCTGATCCCCCTAAGAGTCAACGAGAGGCTGAGAGAGCTTCACTCTCTGTGTCTCAGA 476
DB 883 GCTGATCCCCCTAAGAGTCAACGAGAGGCTGAGAGAGCTTCACTCTCTGTGTCTCAGA 942
QY 477 AATCATGATGATCATCTCTCTGTGTCTTCTTCTACCTCTGTGTGTGTCTCATCGAGATGATA 536
DB |||||
```



CC animals, although no supporting data is given. Suggested activities  
CC include nutritional activity, immune stimulating or suppressing activity,  
CC haematopoiesis regulating activity, tissue growth activity,  
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
CC activity. The EST sequences are also stated to be useful for gene  
XX therapy.

SQ Sequence 471 BP; 93 A; 126 C; 143 G; 108 T; 1 other;

Query Match 65.1%; Score 422; DB 20; Length 471;

Best Local Similarity 99.5%; Pred. No. 1.3e-105;

Matches 433; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGCTGCTCCCTCAAGATGTTTCCCTGCTGCTCTCTCGTGTATCTACTGGTCACT 60  
DB 29 ATGCTGCTCCCTCAAGATGTTTCCCTGCTGCTCTCTCGTGTATCTACTGGTCACT 88  
QY 61 GTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 89 GTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 147  
QY 121 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
DB 148 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 207  
QY 181 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 208 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267  
QY 241 CACGAGAGTGGAGAGCCCTTTTACGGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 268 CACGAGAGTGGAGAGCCCTTTTACGGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 327  
QY 301 CAGGAGTGTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 328 CAGGAGTGTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387  
QY 361 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
DB 388 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447  
QY 421 ATCCCTTAAAGATC 435  
DB 448 ATCCCTTAAAGATC 462

RESULT 12

AS86763

ID AS86763 standard; cDNA; 3531 BP.

XX AC AS86763;

XX AC AS86763;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #22567.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.

OS Homo sapiens.

XX WO200175067-A2.

FN WO200175067-A2.

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

FR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA (HYSE-) HYSEQ INC.

XX

PI

XX

DR

DR

XX

PT

PT

PT

XX

XX

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX AC AAS86762;  
XX DT 13-FEB-2002 (first entry)  
XX DE DNA encoding novel human diagnostic protein #22566.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX OS food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX PN Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX PR 31-MAR-2000; 2000US-0540217.  
XX PR 23-AUG-2000; 2000US-0649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX DR P-P8DB; ABG22575.  
XX XX  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.  
XX PS Claim 1; SEQ ID No 22566; 103pp; English.  
XX XX  
XX CC The invention relates to isolated polynucleotide (I) and  
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX CC and gene mapping, and in recombinant production of (II). The  
XX CC polynucleotides are also used in diagnostics as expressed sequence tags  
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques  
XX CC to restore normal activity of (II) or to treat disease states involving  
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as  
XX CC a food supplement. (II) and its binding partners are useful in medical  
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating  
XX CC disorders involving aberrant protein expression or biological activity.  
XX CC The polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human  
XX CC diagnostic coding sequences of the invention.  
XX CC Not: The sequence data for this patent did not appear in the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pat\_sequences.  
XX SQ Sequence 621 BP; 156 A; 155 C; 175 G; 135 T; 0 other;  
  
Query Match 34.6%; Score 224.4; DB 23; Length 621;  
Best Local Similarity 99.6%; Pred. No. 1.9e-51;  
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 220 ATTACGAGTATCGGAATGGCCACCCAGGAGGTGGAGAGCCCTTTCAGGGCGCTGAG 279  
DB 58 ATTACGAGTATCGGAATGGCCACCCAGGAGGTGGAGAGCCCTTTCAGGGCGCTGAG 117  
  
QY 280 TGGATGGCAGCAAGACCTGAGGAGCGTGCTCATCTGCTCAACGTCACCTGTGAAC 339  
DB 118 TGGATGGCAGCAAGACCTGAGGAGCGTGCTCATCTGCTCAACGTCACCTGTGAAC 177  
  
QY 340 GACTCTGGCCTTACCTGCAATGTGTCCCGGAGTTTGAGTTCAGGGCGCATCGGCC 399  
|||||

DB 178 GACTCTGGCCTTACACCTGCAATGTGTCCCGGAGTTTGAGTTTGAGGGCATCGGCC 237  
QY 400 TTGTGTAAGACGACGGCGGTGATCCCTTAAGAGTACCCAGGAGG 445  
|||||  
DB 238 TTGTGTAAGACGACGGCGGTGATCCCTTAAGAGTACCTAGGAGG 283  
|||||  
  
RESULT 14  
AAC90603  
ID AAC90603 standard; cDNA; 657 BP.  
XX AC AAC90603;  
XX DT 13-MAR-2001 (first entry)  
XX DE Rabbit sodium channel beta1 protein coding sequence.  
XX KW Rat; sodium channel beta3 protein; Alrx94h5; pain; sleep disorder;  
XX KW neurodegenerative disorder; mood disorder; muscle contraction; ss.  
XX OS Oryctolagus cuniculus.  
XX PN WO200069912-A1.  
XX PD 23-NOV-2000.  
XX PF 12-MAY-2000; 2000WO-US13144.  
XX PR 14-MAY-1999; 99US-0134198.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Curtis RAJ;  
XX DR WPI; 2001-122743/13.  
XX XX  
XX PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal  
XX PT root ganglion cDNA library for use in chromosome mapping, forensic  
XX PT medicine, monitoring clinical trials and therapeutics.  
XX PS Disclosure; Fig 5; 145pp; English.  
XX CC The present invention provides the protein and coding sequences of the  
XX CC rat sodium channel beta3 protein, designated Alrx94h5. This protein is  
XX CC involved in the generation of pain and other sensory or perceptual nerve  
XX CC impulses, in the establishment and endurance of mood, neurodegenerative  
XX CC and sleep disorders, and in the control of muscle contraction, including  
XX CC movements such as the heartbeat, digestion and vascular tone. The  
XX CC sequences can be used in predictive medicine, screening and diagnostic  
XX CC assays, and in pharmacogenomics.  
XX SQ Sequence 657 BP; 146 A; 181 C; 213 G; 117 T; 0 other;

Query Match 22.9%; Score 148.2; DB 22; Length 657;  
Best Local Similarity 57.7%; Pred. No. 1.4e-30;  
Matches 333; Conservative 0; Mismatches 223; Indels 21; Gaps 3;  
  
QY 76 TGTGTGGAAGTGCCTCGGAGACGGAGCGCGTGCAGGCAACCCCATGAGCTCGCTGC 135  
DB 61 TGGTGTGAGGTGGACTCGGAGACCGAGCCGTGTACGGATGACCTTCAAAATTTCTGTC 120  
  
QY 136 ATCTCTCTGATGAAGAGAGAGGAGGTGGAGGCCACCCACCGTGTGGAATGTTCTACAGG 195  
DB 121 ATCTCTCTGAAGCGCCGCGAGAGACCGAGCCCTTCACGGAGTGCAGCTTCGCGC 180  
  
QY 196 CCGAGAGCGCGTAAAGATT-----TCCTATTATTCAGTATCGGAATGGCCACAG 246  
DB 181 CAGAAGGCGCATCGAGGAGTTTCGTCAGATCTATGAGAACCGAGGTGCTGCAGCTG 240  
  
QY 247 GAGGTGGAGAGCCCTTTCAGGGCGCGCTGCAGTGAATGGCAG-----CAAGGAC 297  
DB 241 GAGAAGACGAGCGCTTTGAGGGCCCGCTGTGTGGAACCGGAGCGGGGACCAAGGAC 300  
|||||



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:08:07 ; Search time 1915.74 Seconds  
(without alignments)  
8221.015 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_376\_1023

Perfect score: 648

Sequence: 1 atgcctgccttcaatagatt.....cgtaccagtggaaggaatag 648

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	646	99.7	1201	13	BX420015
2	624.4	96.4	1201	13	BX445002
3	536.6	82.8	1359	11	AK076466 Mus muscu
4	536.6	82.8	3549	11	AK049747 Mus muscu

5	525.6	81.1	4149	11	AK049286
6	518	79.9	636	12	BM933157
7	510.8	78.8	672	10	BB614118
8	468	72.2	950	10	BG294174
9	447.6	69.1	723	14	CA749311
10	408.4	63.0	927	13	BQ713131
11	385.2	59.4	825	13	BQ745919
12	382.4	59.0	846	14	CD355879
13	370.4	57.2	1069	12	BM928131
14	357.8	55.2	742	14	CB526211
15	351	54.2	986	13	BUI18914
16	348.4	53.8	845	14	CA327438
17	343.6	53.0	700	12	BI739617
18	336.2	51.9	652	10	BB652801
19	330.6	51.0	892	13	BUI368614
20	326	50.3	595	13	BUI368531
21	310	47.8	450	14	CB787935
22	301.6	46.5	712	14	CB526257
23	299.6	46.2	780	13	BQ770528
24	291.6	45.0	750	14	CD349206
25	273	42.1	669	12	BM951151
26	266.4	41.1	471	14	CB732717
27	258.8	39.9	705	14	CD349958
28	243.6	37.6	604	14	CB581657
29	242.4	37.4	710	10	BB653495
30	221.8	34.2	669	14	CB520894
31	210	32.4	683	13	BU492633
32	189.6	29.3	638	10	BB657468
33	188.6	29.3	1507	11	AK051805
34	179.2	27.7	490	13	BY247892
35	172	26.5	620	10	BB651905
36	171.6	26.5	600	12	BI990184
37	171.6	26.5	600	12	BI990395
38	167	25.8	1201	9	AL534136
39	148	22.8	603	12	BM490673
40	144.4	22.3	319	13	BY129048
41	143.8	22.2	822	12	BI755159
42	143	22.1	1037	12	BM920864
43	142.6	22.0	1119	12	BM805468
44	136.4	21.0	405	10	BE956050
45	134.8	20.8	847	12	BI549123

ALIGNMENTS

RESULT 1	BX420015	BX420015	Homo sapiens	FETAL BRAIN	Homo sapiens	CDNA clone	EST 13-MAY-2003
LOCUS	BX420015	CS9DF023YA09	5-PRIME	MRNA	sequence.		
DEFINITION	BX420015	BX420015	1	GI:30646738			
ACCESSION	BX420015	BX420015	1	GI:30646738			
VERSION	BX420015	BX420015	1	GI:30646738			
KEYWORDS	EST.	EST.					
SOURCE	Homo sapiens	Homo sapiens					
ORGANISM	Homo sapiens	Homo sapiens					
REFERENCE	1	(bases 1 to 1201)					
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.						
TITLE	Full-length cDNA libraries and normalization						
JOURNAL	Unpublished						
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6147.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS9DF023YA05QP1&cluster=6147.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/invitrogen Corporation 1600						









of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 4149)  
 REFERENCE  
 AUTHORS  
 Adachi, J., Aizawa, K., Hanagaki, T., Hara, A., Hagihizumi, W.,  
 Fukuda, S., Furuno, M., Nakamura, T., Miyazaki, A., Murata, M.,  
 Hayashida, K., Hayashi, N., Hiramoto, K., Hiraoka, T., Hirozawa, T.,  
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M., and Hayashizaki, Y.  
 TITLE  
 JOURNAL  
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-3216)  
 COMMENT  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/  
 FEATURES  
 source  
 1. 4149  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:C330019103"  
 /db\_xref="taxon:10090"  
 /cell\_type="ES cells"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 misc\_feature  
 410. 1058  
 /note="VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM  
 CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus]  
 (SPR|09JK00, evidence: FASTV, 99.5%id, 100%length,  
 match=66)  
 putative"  
 4129. 4134  
 /note="putative"  
 polyA\_site  
 4149  
 /note="putative"  
 BASE COUNT  
 1088 a 1022 c 961 g 1077 t 1 others  
 ORIGIN  
 Query Match 81.1%; Score 525.6; DB 11; Length 4149;  
 Best Local Similarity 89.2%; Pred. No. 2.8e-130;  
 Matches 578; Conservative 0; Mismatches 69; Indels 1; Gaps 1;  
 QY 1 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCAGT 60  
 DB 410 ATGCTGCTTCAACAGATTGTTCCCTAGCTTCTCTAGTGTCTACTGGGTCAGA 469  
 QY 61 GTCTGCTTC-CCTGTGTGTGGAGTCCCTCGAGACCGAGCCGTCGAGGCAACCC 119  
 DB 470 GTCTGCTTCNCCTGTGTGTGAGAGTACCTCGGACAGAGCCGTCGAGGCAATTC 529  
 QY 120 CATGAAGCTGGCTGCATCTCTCTGCATGAAGAGAGAGGAGTGGAGCCACACCGTGTG 179  
 DB 530 CATGAAGCTGAGATGCATCTCTCTGATGAAGAGAGGAGGAGTGGAGCCACCTGATG 589  
 QY 180 GGAATGTTCTACAGCCCGAGCGCGGTAAAGATTTCCTATTATACGATATCGGAATGG 239  
 DB 590 GGAGTGGTCTACAGCCCTGAGCGCGGTAAAGATTTCCTATTATATGATATCGAATGG 649

240 CCACACGAGGTGGAGAGCCCTTTACAGGCGCGCTGAGTGGATGSCACAGGACCT 299  
 DB 650 CCACACGAGGTGGAGAGCCCTTTCCAAAGGTGCTGTCAGTGGATGAGGAGCAAGACCT 709  
 QY 300 GCAGGACGCTGTCATCACTGCTCAACGCTCACTCTGAACGACTTCTGGCCTCTACACCTG 359  
 DB 710 GCAGGACGCTATCACTCACTGCTTCTCAATGTCACCTGATGACTCTGGCCTCTACACATG 769  
 QY 360 CAATGTGTCCCGGAGTTGAGTTGAGCGCATCGGCGCTTGTGAGAGAGAGCGCGCT 419  
 DB 770 TAATGTGTCCCGGAGTTGAGTTGAGCGCATCGGCGCTTGTGAGAGAGAGCAACAAGACT 829  
 QY 420 GATCCCGCTTAAGAGTCAACGAGCGCTGAGAGAGACTTCACTCTCTGTGGTCTCAGAAAT 479  
 DB 830 AATACCCCTGCGAGTCACTGAAGAGCGCGGAGAGAGACTTCACTCTCGTGGTCTCGGAAAT 889  
 QY 480 CATGATGATCATCTCTGCTTCTCTCACTCACTCTGGCTGCTCATCGAGATGATATTTG 539  
 DB 890 CATGATGATCATCTCTCTCTCACTCACTCTGGCTGCTTATTTGAGATGATCTATTG 949  
 QY 540 CTACAGAAAGTCTCTCAAAAGCCGAGAGAGCGCCCAAGAAACCGCTCTGACTACCTTGC 599  
 DB 950 CTACAGAAAGTCTCTCAAGCGCGAGAGAGAGCTCAGGAAATGCTCTGACTACCTTGC 1009  
 QY 600 CATCCCATCTGAGACAGAGAGACTCTGCGTACCATGAGGAGGATA 647  
 DB 1010 TATCCCTTCAGAGAACAGAGAGACTCTGTGTACCCCGTGGAGGAATA 1057

RESULT 6  
 LOCUS  
 DEFINITION  
 UI-M-BH3-bag-d-04-0-UI.r1 NIH BMAP\_M\_S4 Mus musculus CDNA clone  
 UI-M-BH3-bag-d-04-0-UI 5', mRNA sequence.  
 ACCESSION  
 BM933157  
 VERSION  
 BM933157.1 GI:19392309  
 EST.  
 KEYWORDS  
 Mus musculus (house mouse)  
 SOURCE  
 Mus musculus  
 ORGANISM  
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT  
 Genom. Res. 6 (9), 791-806 (1996)  
 97044477  
 8889548  
 Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mEstemal.nih.gov  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.regen.com).  
 Seq primer: M13 REVERSE.  
 Location/Qualifiers  
 1. 636  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH3-bag-d-04-0-UI"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NIH BMAP\_M\_S4"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified

	BASE COUNT	158 a	154 c	169 g	155 t	
polylinker; Site 1: Not I; Site 2: Eco RI; The NIH BMAP M <sub>94</sub> library is a subtracted library of a series, ultimately derived from a mixture of individually tagged, normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. the following serially subtracted libraries were generated in this process: NIH BMAP M <sub>94</sub> , NIH BMAP M <sub>93.3</sub> , NIH BMAP M <sub>93.2</sub> , NIH BMAP M <sub>93.1</sub> , NIH BMAP M <sub>92</sub> , NIH BMAP M <sub>91</sub> . The subtracted library (NIH BMAP M <sub>94</sub> ) was constructed as follows: PCR-amplified cDNA inserts from NIH BMAP M <sub>93.3</sub> , NIH BMAP M <sub>93.2</sub> , and NIH BMAP M <sub>93.1</sub> clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH BMAP M <sub>93.3</sub> , NIH BMAP M <sub>93.2</sub> , and NIH BMAP M <sub>93.1</sub> libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH BMAP M <sub>94</sub> library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)." 5791-806, 1996"	Query Match	79.9%	Score 518;	DB 12;	Length 636;	
	Best Local Similarity	88.9%;	Pred. No. 1.2e-128;			
	Mismatches 560;	Conservative 0;	Mismatches 70;	Indels 0;	Gaps 0;	
QY	1	ATGCCTCGCTCAATAGATTGTTCCCTCGCGCTCTCTCGCGCTTACTCTACTGGTTCAGT	60			
Db	7	ATGCTTCGCTTCAACAGTTGCTTCCCTAGATTCTTAGTGCTCACTACTGGGTGAGA	66			
QY	61	GTCTGCTTCCCTGTGTGTGTGGAAATGCCCTCGAGACGAGAGCGCGCTGCAGGGCAACCCC	120			
Db	67	GTCTGCTTCCCTGTGTGTGTAGAAATACCTCGAGACAGAAAGCCGTGCAGGGCAATTC	126			
QY	121	ATGAAGCTGCGCTGCATCTCCTCATGAAGAGAGAGAGGTGAGAGCCACCAAGCTGGTG	180			
Db	127	ATGAAGCTGAGATGCATCTCTCATGATAGAGAGGAGGAGGTGAAGCCACCACTGTATG	186			
QY	181	GAATGGTTCATACGGCCCGAGGGCGGTAAAGATTTCTTATTACGAGTATCGGAATGGC	240			
Db	187	GAGTGGTTCATACGGCTGAGGGCGGTAAAGATTTCTTATATGATGATATCGAAATGGC	246			
QY	241	CACAGGAGGTGGAGAGCCCTTTACGGGGCGCTGCGAGTGGAAATGGCAGCAAGGACCTG	300			
Db	247	CACAGGAGGTGGAGAGCCCTTCCAAAGTGTCTGCAATGGGAGGAGCAAGACTG	306			
QY	301	CAGGAGTGTCCATCACTGTGCTCAACGTCACTCTGAAGCACTTGGGCTCTACACCTGC	360			
Db	307	CAGGAGTGTCCATCACTGTGCTCAACGTCACTCTGAATGATCTTGGGCTCTACACATGT	366			
QY	361	AATGTGTCGGGAGTTTGAGTTTGAGGGCGCATCGCGCTTTGTGAAGACGACGCGCTG	420			
Db	367	AATGTGCCAGGAGTTTGAGTTTGAAGCACACCGGCCCTTTGTGAAGACCAAGACTA	426			
QY	421	ATCCCCCTAAGATCACCGAGAGGCTGGAGAGCACTCACTCTGTGGTCTCAGAATC	480			
Db	427	ATACCCCTGGAGTCACTGAAGAGGGGGAGAGAGACTTCACTCCGTGGTCTCGGNAATC	486			
QY	481	ATGATGTACATCTTCTGGTCTTCCCTCAACCTGTGGCTGCTCATCGAGATGATATTGC	540			
Db	487	ATGATGTACATCTTCTGGTCTTCCCTCACTCTGTGGCTGTTTATTGAGATGATCTATTGC	546			
QY	541	TACAGAAAGTCTCABAAGCCGAGAGGAGCCCAAGAAACCGCTCTGACTACTCTGCC	600			
Db	547	TACAGAAAGTCTCTAAGGTCGAGAGGAGGAGCTCAGAGAAATGGGTCTGCTTACCTTGC	606			
QY	601	ATCCCATCTGAGAAACAAGGAGAACTCTGGC	630			

/note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCCAAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTAAATAATCCCCCCCCCC sequence 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I." 170 a 167 c 172 g 163 t

BASE COUNT 170 a 167 c 172 g 163 t  
 ORIGIN  
 Query Match 78.8%; Score 510.8; DB 10; Length 672;  
 Best Local Similarity 89.2%; Pred. No. 1.1e-126;  
 Matches 551; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
 QY 1 ATCCCTGCTTCAATAGATGTTCCCTGGCTCTCTCGTCTATCTACTGCTGCTAGT 60  
 DB 55 ATCCCTGCTTCAACAGATGCTTCCCTAGTCTCTCTAGTCTATCTACTGCTGCTAG 114  
 QY 61 GTCTGCTTCCCTGCTGTGTGAAGTGCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 DB 115 GTCTGCTTCCCTGCTGTGTGAGAGTACCTCGAGAGAGAGAGAGAGAGAGAGAGAG 174  
 QY 121 ATGAGCTGCTGCTATCTCTGTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 DB 175 ATGAGCTGAGATGCTATCTCTGTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 234  
 QY 181 GAATGGTTCACAGGCGCGAGGCGGTAAAGATTTCTTATTTAGAGTATCGGAATGGC 240  
 DB 235 GAGTGGTTCACAGGCGCTGAGGCGGTAAAGATTTCTTATTTAGATATGATATCGGA 294  
 QY 241 CACAGAGAGTGAAG 300  
 DB 295 CACAGAGAGTGAAG 354  
 QY 301 CAGAGAGTGTCTCACTGCTCTACAGTCACTCTGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 DB 355 CAGAGAGTATCCATCACTGCTCTCAATGTCACCTGATGACTCTGAGCTCTACAGAGT 414  
 QY 361 AATGTGTCGGAGAGTTCAGTTGAGGCGGATCGCCCTTGTGCAAGAGAGAGAGAGAG 420  
 DB 415 AATGTGTCAGGAGAGTTCAGTTGAGGCGGATCGCCCTTGTGCAAGAGAGAGAGAGAG 474  
 QY 421 ATCCCTTAAGAGTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 DB 475 ATACCCCTGCGAGTCACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534  
 QY 481 ATGATGATATCTTCTGCTCTCTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 DB 535 ATGATGATATCTTCTGCTCTCTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594  
 QY 541 TACAGAAAGTCTCAAAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 DB 595 TACAGAAAGTCTTAAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654  
 QY 601 ATCCCTTCTGAGAGAGAG 618  
 DB 655 ATCCCTTCTGAGAGAGAG 672

RESULT 8  
 LOCUS BG294174 950 bp mRNA linear EST 21-FEB-2001  
 DEFINITION 602391245F1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:4503250 5',

ACCESSION BG294174  
 VERSION BG294174.1 GI:13054543  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

mRNA sequence.

1 (bases 1 to 950)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov

Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10373 row: 1 column: 11  
 High quality sequence stop: 643.

FEATURES  
 Location/Qualifiers  
 1..950  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4503250"  
 /tissue\_type="retina"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_94"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: Sall; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 3.3 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 230 a 257 c 260 g 203 t

Query Match 72.2%; Score 468; DB 10; Length 950;  
 Best Local Similarity 88.8%; Pred. No. 4.4e-115;  
 Matches 529; Conservative 0; Mismatches 65; Indels 2; Gaps 2;

QY 1 ATGCTGCTGCTTCAATAGATGTTTCCCTGGCTCTCTCGTCTGCTTATCTACTGCTGCT 60  
 DB 87 ATGCTGCTGCTTCAACAGATGCTTCCCTAGTCTCTCTAGTCTCTACTGCTGCTGCT 146  
 QY 61 GTCTGCTTCCCTGCTGTGTGTGAGTGCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 DB 147 GTCTGCTTCCCTGCTGTGTGTGAGTGCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAG 206  
 QY 121 ATGAGAGTGCCTGCTGCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 DB 207 ATGAGAGTGCCTGCTGCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 266  
 QY 181 GAATGCTTCAAG 240  
 DB 267 GAGTGGTTCACAG 326  
 QY 241 CACAGAGAGTGCAG 300  
 DB 327 CACAGAGAGTGCAG 386  
 QY 301 CAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 DB 387 CAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446  
 QY 361 AATGTGTCGGAGAGTTCAGTTCAGGCGGATCGCCCTTGTGAGAGAGAGAGAGAGAG 420  
 DB 447 AATGTGTCGGAGAGTTCAGTTCAGGCGGATCGCCCTTGTGAGAGAGAGAGAGAGAG 505  
 QY 421 ATCCCTTCTGAG 480

Db 506 ATACCCCTCGAGTCACTGAAGAGCGGAGAGAGACTTCACTCCGCTGCTCGGAATC 565  
 Qy 481 ATGATGATACATCTTCTGCTCTCTCCACCTGCTGCTCATCGAGATGATATATTC 540  
 Db 566 ATGATGATACATCTTCTGCTCTCTCCACCTGCTGCTCATCGAGATGATATATTC 625  
 Qy 541 TACAGAAAGTCTTCAAAAGCCGAGAGAGCGAGCCCAAGAAAGCGCTCTGACTACT 596  
 Db 626 TACAGAAAGTCTTCAAGAGCGGAGAGAGCGAG-TGAGAAATGCGCTGACTACTT 680

RESULT 9  
 CA749311 723 bp mRNA linear EST 27-NOV-2002  
 LOCUS UI-M-FY0-cdd-k-17-0-UI.r1 NIH\_BMAP\_FY0 Mus musculus cDNA clone  
 DEFINITION IMAGE: 6831594 5', mRNA sequence.  
 CA749311  
 CA749311.1 GI:25570984  
 EST.  
 Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 723)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

Seq primer: pYX-5.  
 Location/Qualifiers  
 1..723  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE: 6831594"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_FY0"  
 /notes="Organ: Brain; Vector: pYX-Asc; Site: 1; EcoR I;  
 Site 2: Not I; The library was constructed according  
 to the protocol of Benton and Davis (1978). The library was  
 screened by hybridization to a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is AGGAGACAG. This library was created for the University  
 of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator."

BASE COUNT 176 a 179 c 201 g 165 t 2 others  
 ORIGIN

Query Match 69.1%; Score 447.6; DB 14; Length 723;  
 Best Local Similarity 89.2%; Pred. No. 1.2e-109;  
 Matches 494; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

Qy 1 ATGCTGCTTCAATAGATGTTTCCCTGCTTCTCTGCTTATCTACTGGTCACT 60  
 Db 170 ATGCTGCTTCAATAGATGTTTCCCTGCTTCTCTGCTTATCTACTGGTCACT 229  
 Qy 61 GTCTGCTTCCCTGCTTGTGTGAAGTGCCTCGAGAGCGGAGCGGTCGAGGCAACCC 120  
 Db 230 GTCTGCTTCCCTGCTTGTGTGAAGTGCCTCGAGAGCGGAGCGGTCGAGGCAATTC 289  
 Qy 121 ATGAAGCTCGCGTGCATCTCTGATGAAGAGAGAGAGAGTGGAGGCCACACCGTGTG 180  
 Db 290 ATGAAGCTGAGATGATCTCTGATGAAGAGAGAGAGAGTGGAGGCCACACCGTGTG 349  
 Qy 181 GAATGCTTCTACAGGCCCGAGGCGGTAAAGATTTCTTATTTACAGATATCGGAATGCC 240  
 Db 350 GAGTGTGTTCTACAGGCCCGTGAAGGCGGTAAAGATTTCTTATTTACAGATATCGGAATGCC 409  
 Qy 241 CACCAAGAGTGGAGAGCGCCCTTTTCAGGGCGCTGAGTGAAGTGGAGAGAGTGGAGAGTGG 300  
 Db 410 CACCAAGAGTGGAGAGCGCCCTTTTCAGGGCGCTGAGTGAAGTGGAGAGAGTGGAGAGTGG 469  
 Qy 301 CAGGACGTGCTCATCACTGCTGCTCAAGTCTCTGAGAGTCTGCGCTCTACACCTGC 360  
 Db 470 CAGGACGTATCCATCACTGTTCTCAATGCTCACTGAGTCTGCGCTCTACACATGT 529  
 Qy 361 AATGTGTCCTCGGAGTGTGAGTTTGAAGCGCATCGGCTTGTGAAGAGAGAGCGGCTG 420  
 Db 530 AATGTGTCCTCGGAGTGTGAGTTTGAAGCGCATCGGCTTGTGAAGAGAGAGCGGCTG 589  
 Qy 421 ATCCCTTAAAGAGTCAACCG-AGGAGGCTGGAGAGAGTCACTCTCTGTGTTCTCAGAAAT 479  
 Db 590 ATACCTTAAAGAGTCACTCAAGAGAGCGGAGAGAGTCACTCTCTGTGTTCTCAGAAAT 649  
 Qy 480 CATGATGATACATCTTCTGCTTCTTCTCACCCTGCTGCTCATCGAGATGATATATG 539  
 Db 650 CATGATGATACATCTTCTGCTTCTTCTCACCCTGCTGCTCATCGAGATGATATATG 709  
 Qy 540 CTACAGAAAGTCT 553  
 Db 710 CTACAGAAAGTCT 723

RESULT 10  
 BQ713131  
 LOCUS BQ713131  
 DEFINITION 927 bp mRNA linear EST 16-JUL-2002  
 5', mRNA sequence.  
 BQ713131  
 BQ713131.1 GI:21852030  
 EST.  
 Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 927)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 Tissue Procurement: Susan L. Sullivan, Ph.D.  
 cDNA Library preparation: Resgen, Invitrogen Corp  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13727 row: b column: 23  
 High quality sequence stop: 679.

FEATURES  
 source  
 1..927  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:21:16 ; Search time 57.0404 Seconds  
(without alignments)  
5014.271 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_376\_1023

Perfect score: 648  
Sequence: 1 atcgctgccttcacatgatt.....cggaccaccagtgaggaatag 648

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.4	5.6	1668	3	US-09-385-028-20
2	36.4	5.6	1668	4	US-09-726-614-20
3	36.4	5.6	1604	3	US-09-385-028-13
4	36.4	5.6	1604	4	US-09-726-614-13
5	36.4	5.6	15079	3	US-09-385-028-1
6	36.4	5.6	15079	4	US-09-726-614-1
7	36.2	5.6	1371	4	US-09-996-243-388
8	35.8	5.5	1236	4	US-09-423-439-9
9	35.8	5.5	1317	4	US-09-252-991A-3363
10	35.8	5.5	1758	4	US-09-252-991A-3311
11	35.8	5.5	1926	4	US-09-423-439-25
12	35.8	5.5	1929	4	US-09-423-439-15
13	35.8	5.5	1995	4	US-09-252-991A-3322
14	35.8	5.5	1998	4	US-09-423-439-50
15	35.8	5.5	2019	4	US-09-423-439-31
16	35.8	5.5	2025	4	US-09-423-439-37
17	35.8	5.5	2217	4	US-09-423-439-52
18	35.6	5.5	1851	4	US-09-252-991A-2593
19	35.6	5.5	2172	4	US-09-252-991A-2325
20	35.4	5.5	2769	3	US-09-118-408-1
21	35.4	5.5	2769	4	US-09-506-855-1
22	35.4	5.5	2769	4	US-09-911-176B-1
23	35.4	5.5	2769	4	US-09-619-740-1
24	35.4	5.5	2769	4	US-09-506-852-1
25	35.2	5.4	894	4	US-09-252-991A-2512
26	35.2	5.4	2681	1	US-08-070-165F-9
27	35.2	5.4	2681	2	US-08-885-418-9

c 28	35.2	5.4	10763	1	US-08-761-258-1	Sequence 1, Appli
c 29	35.2	5.4	10763	2	US-08-977-306-1	Sequence 1, Appli
c 30	34.8	5.4	1296	4	US-09-352-991A-16548	Sequence 16548, A
c 31	34.8	5.4	1689	4	US-09-352-991A-16054	Sequence 16054, A
c 32	34.6	5.4	1724	4	US-09-117-860-36	Sequence 36, Appli
c 33	34.6	5.3	867	4	US-09-252-991A-9371	Sequence 9371, Ap
c 34	34.6	5.3	1296	4	US-09-252-991A-9398	Sequence 9398, Ap
c 35	34.6	5.3	3843	4	US-09-252-991A-1430	Sequence 1430, Ap
c 36	34.6	5.3	5337	4	US-09-252-991A-1588	Sequence 1588, Ap
c 37	34.4	5.3	645	4	US-09-252-991A-8559	Sequence 8559, Ap
c 38	34.4	5.3	1539	4	US-09-252-991A-8610	Sequence 8610, Ap
c 39	34.2	5.3	311	4	US-09-702-705-367	Sequence 367, App
c 40	34.2	5.3	311	4	US-09-736-457-367	Sequence 2, Appli
c 41	34	5.2	1519	1	US-08-225-477B-2	Sequence 2, Appli
c 42	34	5.2	1519	5	PCT-US95-04353-2	Sequence 151, App
c 43	33.8	5.2	683	2	US-08-997-060-151	Sequence 151, App
c 44	33.8	5.2	683	2	US-08-997-362-151	Sequence 151, App
c 45	33.8	5.2	683	3	US-09-095-855-151	Sequence 151, App

## ALIGNMENTS

### RESULT 1

US-09-385-028-20  
; Sequence 20, Application US/09385028  
; Patent No. 6232106  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Parakkar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6232106  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN PPLC  
; STREET: The Jenner Building, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09385,028  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 1418/P57452US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 39305350  
; TELEX: RCA 248593 IDEA UR  
; INFORMATION FOR SEQ ID NO. 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1668 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-385-028-20

Query Match 5.6%; Score 36.4; DB 3; Length 1668;  
Best Local Similarity 47.4%; Pred. No. 1.3;

Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 243 CCAGGAGGTGGAGAGCCCTTTTCAGGGCGCCCTGCGAGTGAATGGCAGGACCTGCA 302  
|||  
Db 1050 CCTCGACGGCTACAAGCACTTCGACCGCTACCGGTCGGCCCGGAGCGACCT 1109  
|||  
QY 303 GGACGTGTCCATCACTGTGCTCAACGTCCTCTGAACGACTCTGAGCTCTACACCTGCAA 362  
|||  
Db 1110 GGAGGCGCGCGCGAGCTGAAGCTGCGCGGGATGCCCGAGCGCTTCGCGACAGGAT 1169  
|||  
QY 363 TGTGTCCCGGAGTTGAGTTGAGCGCATCGGCCCTTTGTGAAGACGACGCGGTGAT 422  
|||  
Db 1170 CGCGCGCGCGAGAGCGGCTCAAGGAGTACCGCGCGCGCGAGCGCTGCGCGCGGCT 1229  
|||  
QY 423 CCCCTTAAGAGTCACCGAGGAGCTGAGAGGACTTCACCTCTGTGGTCT 472  
|||  
Db 1230 CGCGCGGTGCGCATCGAGCGGAGGTGCTGGACTTCCTCGCGGCGACT 1279  
|||

## RESULT 2

US-09-726-614-20  
; Sequence 20, Application US/09726614  
; Patent No. 6514735  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Paradkar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6514735  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
; STREET: The Jennifer Building, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/726,614  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 1418/PS7452US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 39305350  
; TELEX: RCA 248593 IDEA UR  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1668 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)

US-09-726-614-20

Query Match 5.6%; Score 36.4; DB 4; Length 1668;  
Best Local Similarity 47.4%; Pred. No. 1.3; Indels 0; Gaps 0;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 243 CCAGGAGGTGGAGAGCCCTTTTCAGGGCGCCCTGCGAGTGAATGGCAGGACCTGCA 302  
|||  
Db 1050 CCTCGACGGCTACAAGCACTTCGACCGCTACCGGTCGGCCCGGAGCGACCT 1109  
|||  
QY 303 GGACGTGTCCATCACTGTGCTCAACGTCCTCTGAACGACTCTGAGCTCTACACCTGCAA 362  
|||  
Db 1110 GGAGGCGCGCGCGAGCTGAAGCTGCGCGGGATGCCCGAGCGCTTCGCGACAGGAT 1169  
|||  
QY 363 TGTGTCCCGGAGTTGAGTTGAGCGCATCGGCCCTTTGTGAAGACGACGCGGTGAT 422  
|||  
Db 1170 CGCGCGCGCGAGAGCGGCTCAAGGAGTACCGCGCGCGCGAGCGCTGCGCGCGGCT 1229  
|||  
QY 423 CCCCTTAAGAGTCACCGAGGAGCTGAGAGGACTTCACCTCTGTGGTCT 472  
|||  
Db 1230 CGCGCGGTGCGCATCGAGCGGAGGTGCTGGACTTCCTCGCGGCGACT 1279  
|||

Db 1050 CCTCGACGGCTACAAGCACTTCGACCGCTACCGGTCGGCCCGGAGCGACCT 1109  
QY 303 GGACGTGTCCATCACTGTGCTCAACGTCCTCTGAACGACTCTGAGCTCTACACCTGCAA 362  
|||  
Db 1110 GGAGGCGCGCGCGAGCTGAAGCTGCGCGGGATGCCCGAGCGCTTCGCGACAGGAT 1169  
|||  
QY 363 TGTGTCCCGGAGTTGAGTTGAGCGCATCGGCCCTTTGTGAAGACGACGCGGTGAT 422  
|||  
Db 1170 CGCGCGCGCGAGAGCGGCTCAAGGAGTACCGCGCGCGCGAGCGCTGCGCGCGGCT 1229  
|||  
QY 423 CCCCTTAAGAGTCACCGAGGAGCTGAGAGGACTTCACCTCTGTGGTCT 472  
|||  
Db 1230 CGCGCGGTGCGCATCGAGCGGAGGTGCTGGACTTCCTCGCGGCGACT 1279  
|||

## RESULT 3

US-09-385-028-13  
; Sequence 13, Application US/09385028  
; Patent No. 6232106  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Paradkar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6232106  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
; STREET: The Jennifer Building, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/385,028  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 1418/PS7452US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 39305350  
; TELEX: RCA 248593 IDEA UR  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11604 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)

US-09-385-028-13

Query Match 5.6%; Score 36.4; DB 3; Length 11604;  
Best Local Similarity 47.4%; Pred. No. 2.7;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 243 CCAGGAGGTGGAGAGCCCTTTTCAGGGCGCCCTGCGAGTGAATGGCAGGAGGACCTGCA 302  
|||  
Db 8258 CCTCGACGGCTACAAGCACTTCGACCGCTACCGGTCGGCCCGGAGCGCTACCGGCGGCT 8317  
|||  
QY 303 GGACGTGTCCATCACTGTGCTCAACGTCCTCTGAACGACTCTGAGCTCTACACCTGCAA 362  
|||  
Db 8258 CCTCGACGGCTACAAGCACTTCGACCGCTACCGGTCGGCCCGGAGCGCTACCGGCGGCT 8317  
|||  
QY 303 GGACGTGTCCATCACTGTGCTCAACGTCCTCTGAACGACTCTGAGCTCTACACCTGCAA 362  
|||  
Db 8258 CCTCGACGGCTACAAGCACTTCGACCGCTACCGGTCGGCCCGGAGCGCTACCGGCGGCT 8317  
|||

Db 8318 GGAGGCGCCGCGCGAGCTGAAGCTGCGCGGGATGCCGAGCGCTCCGACACAGGAT 8377  
QY 363 TGTGTCCCGGAGTTTGAGTTTGAGCGGCATCGGCCCTTTGTGAAGACGACGCGGCTGAT 422  
Db 8378 CGCCGCGCGAGGACCGGCTCAAGAGTACCGGCGCCGCGAGGCGCTGGCGCGGCT 8437  
QY 423 CCCCCTAAGAGTACCGGAGGCGCTGAGAGGACTTCACCTCTGTGGTCT 472  
Db 8438 CGCCCGGCTCGGCATCGAGCGGCGAGGTGCTGGAATTCCTCGGCGGACT 8487

## RESULT 4

US-09-726-614-13  
; Sequence 13, Application US/09726614  
; Patent No. 6514735  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Paradkar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6514735  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN PPLC  
; STREET: The Jenifer Building, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/726,614  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 1418/P57452US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 39305350  
; TELEX: RCA 248593 IDEA UR  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11604 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-09-726-614-13

Query Match 5.6%; Score 36.4; DB 4; Length 11604;  
Best Local Similarity 47.4%; Pred. No. 2.7;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 243 CCAGAGGTGAGAGCGCCCTTTTCAGGGCGCCTGCACTGCAAGTGGCGAGGACCTGCA 302  
Db 8258 CCTCGAGGCTTACAGACACTTCGACCGTACCGGTCGGCCCGAGGCGACCGGACCT 8317  
QY 303 GGAGGTCTCATCTGCTGCTCAAGCTCTGAGAGCTCTGGGCTCTACACCTGCAA 362  
Db 8318 GGAGGCGCGCGCGCGAGCTGAGCTGGCGGATGCCGAGCGCTTCGCGACAGGAT 8377  
QY 363 TGTGTCCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGAT 422

Db 8378 CGCCGCGCGCAAGACCGGCTCAAGAGTACCGGCGCCGCGAGCGCTGGCGCGGCT 8437  
QY 423 CCCCCTAAGAGTACCGGAGGCGCTGAGAGGACTTCACCTCTGTGGTCT 472  
Db 8438 CGCCCGGCTCGGCATCGAGCGGCGAGGTGCTGGAATTCCTCGGCGGACT 8487  
RESULT 5  
US-09-385-028-1  
; Sequence 1, Application US/09385028  
; Patent No. 6232106  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Paradkar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6232106  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN PPLC  
; STREET: The Jenifer Building, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/385,028  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 1418/P57452US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 39305350  
; TELEX: RCA 248593 IDEA UR  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15079 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptomyces clavuligerus  
; US-09-385-028-1

Query Match 5.6%; Score 36.4; DB 3; Length 15079;  
Best Local Similarity 47.4%; Pred. No. 3;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 243 CCAGAGGTGAGAGCGCCCTTTTCAGGGCGCCTGCAATGGCGAGGACCTGCA 302  
Db 10290 CCTCGAGGCTTACAGACACTTCGACCGTACCGGTCGGCCCGAGGCGACCGGACCT 10349  
QY 303 GGAGGTCTCATCTGCTGCTCAAGCTCTGAAACGACTCTGGGCTCTACACCTGCAA 362  
Db 10350 GGAGGCGCGCGCGCGAGCTGAAGCTGGCGGATGCCGAGCGCTTCGCGACAGGAT 10409  
QY 363 TGTGTCCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGAT 422  
Db 10410 CGCCGCGCGCAAGAGCGCGCTCAAGGAGTACCGGCGCGGCGGCTGGCGCGGCT 10469

QY 423 CCCCTAAGAGTACCCAGAGAGGCTGGAGAGGACTTCACTCCTCTGTGGTCT 472  
DB 10470 CGCCCGGGTCGGCATCGAGCGGAGGTGCTGGACTTCCCGTCGGCGGACT 10519

## RESULT 6

US-09-726-614-1  
; Sequence 1, Application US/09726614  
; Patent No. 6514735  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Padkar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6514735  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
; STREET: The Jenifer Building, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA: US/09/726,614  
; APPLICATION NUMBER: US/09/726,614  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 1418/P57452U52  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 39305350  
; TELEX: RCA 248593 IDEA UR  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15079 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptomyces clavuligerus  
US-09-726-614-1

Query Match 5.6%; Score 36.4; DB 4; Length 15079;  
Best Local Similarity 47.4%; Pred. No. 3;  
Matches 109; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
QY 243 CCAGAGGTGAGAGCCCTTTACGGCGGCTCAGTGGNATGGCAGAGAGCTGCA 302  
DB 10290 CUTCAGCGGTACAGAGCACTTCGACCGGTACCGGCTCGGCGCCGAGGSCACCGCGGACCT 10349  
QY 303 GGACCTGTCCATCACTGTGCTCAACGTCATCTGACGACTCTCGGCTCTACACCTGCA 362  
DB 10350 GGAGCGCGCCGCGCGAGCTGAGCTGCGCGGGATGCCGACGGCTTCGCGACCAAGAT 10409  
QY 363 TGTGTCCCGGAGTTGATTTGAGGCCCATCGGCGCTTTGTGAAGACGACGCGGCTGAT 422  
DB 10410 CGCCGCGCGAGGAGCGGCTCAAGAGTACCGGCGCGCGCGGCGGCTGCGCGCGGCT 10469

## RESULT 7

US-09-936-243-388  
; Sequence 388, Application US/09996243  
; Patent No. 6478325  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Deanovs, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary B.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C13  
; CURRENT APPLICATION NUMBER: US/09/996,243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04

;;  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252

;;  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 5.6%; Score 36.2; DB 4; Length 1371;

Best Local Similarity 49.2%; Pred. No. 1.4;

Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 198 CGAGGGCGGTAAAGATTCCTTTACGAGTATCGGATGCCACACGAGGTCGAG 257

Db 327 CGGGGACCTGAGCAGTTTGTATCTACCACTAGATCCCTTCCACCATGAGTGG 386

Qy 258 CCCTTTTCAGGGCGCGCTGCAGTGGGATGGCAGCAAGGACCTGCAGGACGTGTCCATCAC 317

Db 387 GCGGTTTAAGGACCGGGTGTCTTGGGATGCGGATCCTGAGCGGTACGATGCTCCATCCT 446



Db 1031 TGTCGATCCCGACCCCTTCAGCGCGCCAGCGGTTTCGGCTACAGGTGGCGGCCGACGTTCCGCC 972  
Qy 245 AGGAGGTGGAGAGCCCTTTTCAGGGCGGCTGAGTGAATGGCAGCAAGGACCTGTCAGG 304  
Db 971 ACAGAGTGTCTCCGACCGCGCGGCTGCTGCGGTTCCACCATCCGACCACTGTTGAAGG 912  
Qy 305 ACCTGTCCATCATCTGCTCAGGTCACTCTGACGACTCTGGCTCTACACCTGCAATG 364  
Db 911 AGCTGTGTGCGGAGCTTCGCGCACCTCGTGATTCGCGGTGAGCTGCACACGCCAAG 852  
Qy 365 TGTCGCGGAGTTTGAGTTGAGGCGCCATCG 395  
Db 851 TGTCGCGGAGACCTGCTTTCACCCATCG 821

## RESULT 11

US-09-423-439-25  
; Sequence 25, Application US/09423439  
; Patent No. 6339070  
; GENERAL INFORMATION:  
; APPLICANT: EMERY, Stephen Charles  
; TITLE OF INVENTION: CHEMICAL COMPOUNDS  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pillsbury Winthrop, L.L.P.  
; STREET: 1100 New York Ave., N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MS Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/423,439  
; FILING DATE: 09-NO. 6339070-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB98/01294  
; FILING DATE: 05-MAY-1998  
; APPLICATION NUMBER: GB 9709421.3  
; FILING DATE: 10-MAY-1997  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1926 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-423-439-25  
Query Match 5.5%; Score 35.8; DB 4; Length 1926;  
Best Local Similarity 52.3%; Pred. No. 2;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
Qy 270 GCGCTTCAGTGAATGGCAGCAAGGACCTGCGAGGAGCTGTCATCATCTGCTCAAGT 329  
Db 1464 GCGCTTCAGTGAATGGCAGCAAGGACCTGCGAGGAGCTGTCATCATCTGCTCAAGT 1523  
Qy 330 CACTCTGAAGACTTGGCTCTACACCTGCATGTGTCGCGGAGTTTGTAGTTTGAGGC 389  
Db 1524 CAGCTGAAGCGCGGAGCTGCGGTACGCGGCAACGAGGACTTCGAGCGCGCATGAGAC 1583  
Qy 390 GCATCGGCCCTTTGTGAAGACGACGCGGCTG 420  
Db 1584 GCTGGAAGACGCGCGCAGCAGAGAGCTG 1614

## RESULT 12

US-09-423-439-15  
; Sequence 15, Application US/09423439  
; Patent No. 6339070  
; GENERAL INFORMATION:  
; APPLICANT: EMERY, Stephen Charles  
; TITLE OF INVENTION: CHEMICAL COMPOUNDS  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pillsbury Winthrop, L.L.P.  
; STREET: 1100 New York Ave., N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MS Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/423,439  
; FILING DATE: 09-NO. 6339070-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB98/01294  
; FILING DATE: 05-MAY-1998  
; APPLICATION NUMBER: GB 9709421.3  
; FILING DATE: 10-MAY-1997  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1929 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:

## US-09-423-439-15

Query Match 5.5%; Score 35.8; DB 4; Length 1929;  
Best Local Similarity 52.3%; Pred. No. 2;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
Qy 270 GCGCTTCAGTGAATGGCAGCAAGGACCTGCGAGGAGCTGTCATCATCTGCTCAAGT 329  
Db 1467 GCGCTTCAGTGAATGGCAGCAAGGACCTGCGAGGAGCTGTCATCATCTGCTCAAGT 1526  
Qy 330 CACTCTGAAGACTTGGCTCTACACCTGCATGTGTCGCGGAGTTTGTAGTTTGAGGC 389  
Db 1527 CAGCTGAAGCGCGGAGCTGCGGTACGCGGCAACGAGGACTTCGAGCGCGCATGAGAC 1586  
Qy 390 GCATCGGCCCTTTGTGAAGACGACGCGGCTG 420  
Db 1587 GCTGGAAGACGCGCGCAGCAGAGAGCTG 1617

## RESULT 13

US-09-252-991A-3322  
; Sequence 3322, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 3322







181	GAATGGTTCTTACAGCCCGAGGGCGGTAAAGATTTTCTCTTATTTCAGAGTATCGGAATGGC	240
556	GAATGGTTCTTACAGCCCGAGGGCGGTAAAGATTTTCTCTTATTTCAGAGTATCGGAATGGC	615
241	CACAGGAGGTGGAGAGGCCCTTTTCAGSGGCGCTGCAGTGGAAATGCGCAGCAGACACTGC	300
616	CACAGGAGGTGGAGAGGCCCTTTTCAGSGGCGCTGCAGTGGAAATGCGCAGCAGACACTGC	675
301	CAGGACGTTCCATCACTGTGCTCAACGTCACCTCTGTGAACGACTCTGGCTCTACACCTGCG	360
676	CAGGACGTTCCATCACTGTGCTCAACGTCACCTCTGTGAACGACTCTGGCTCTACACCTGCG	735
361	AATGTGTCCTCCGGAGTTTGAGTTTGAAGCGCGCATCGGCCCTTTCTGTAAGACGACGCGGCTG	420
736	AATGTGTCCTCCGGAGTTTGAGTTTGAAGCGCGCATCGGCCCTTTCTGTAAGACGACGCGGCTG	795
421	ATCCCCCTAAGAGTCAACGAGAGGCTGGAGAGACTTCACTCTCTGTGTTCTCAGAAATC	480
796	ATCCCCCTAAGAGTCAACGAGAGGCTGGAGAGACTTCACTCTCTGTGTTCTCAGAAATC	855
481	ATGATGTACATCCTTCGTCTTCTCAACCTGTGGCTGCTCATCCAGATGATATATATTCG	540
856	ATGATGTACATCCTTCGTCTTCTCAACCTGTGGCTGCTCATCCAGATGATATATATTCG	915
541	TACAGAAAGGTTCTAAAGCCGAAAGAGGACGCCCAAGAAAAACCGCTCTCACTACCTTGGCC	600
916	TACAGAAAGGTTCTAAAGCCGAAAGAGGACGCCCAAGAAAAACCGCTCTCACTACCTTGGCC	975
601	ATCCCATCTGAGAACAGAGGAGTCTTGGCGGTACCACTGAGGAAATAG	648
976	ATCCCATCTGAGAACAGAGGAGTCTTGGCGGTACCACTGAGGAAATAG	1033

```

RESULT 2
US-09-997-579-3
; Sequence 3, Application US/09997579
; Patent No. US20020113203A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sod
; TITLE OF INVENTION: channel
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses t
; FILE REFERENCES: 674558-2001
; CURRENT APPLICATION NUMBER: US/09/997,579
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: rat
US-09-997-579-3

```

181	QY	GAATGGTCTACAGCCCGGAGGGCGGTAAAGATTCTCTTATTTACGAGTATCGGAATGCG	240
182	QY		
183	QY		
543	Db	GAGTGGTTCTACAGCGCTGAGGGCGGTAAAGATTCTCTTATATGAGTATCGGAATGCG	602
544	Db		
241	QY	CACCAAGGAGTGGAGAGACCCCTTTACAGGGCGCGCTCGAGTGGAAATGGCAGCAAGACCTG	300
242	QY		
603	Db	CACCGAGAGTGGAGAGCCCTTCCAAAGCCGCTGCGAGTGGAAATGGGNGCAGAGCCTG	662
604	Db		
301	QY	CAGGACGTGTCCATCATCTGTGCTCAAGCTCACTCTGAAACGACTCTGSCCTCTACACCTGC	360
302	QY		
663	Db	CAGGACGTATCCATCACTACTCAATGTCACTTTGAAATGACTCTGSCCTCTACACATGC	722
664	Db		
361	QY	TAATGTGTCCTCCGGAGTTTGAGTTTGAGGGCATCGGCCCTTTGTGAAGACGACGCGGCTG	420
362	QY		
723	Db	TAATGTGTCCAGGGAGTTTGGAAATTCAGGGCCACAGGCCCTTTGTGAAGACCAAGAGACTG	782
724	Db		
421	QY	ATCCCCCTAAGAGTACCCGAGGAGGCTGGAGAGGACTTCACCTCTGTGTCTCGAAATC	480
422	QY		
783	Db	ATACCTTTTGCAGTCACTGAAGAGGGCGGAGAGACTTCACCTCCGTGTCTCGGAATC	842
784	Db		
481	QY	ATGATGTACATCTTCTGGTCTTCTCACCCTGTGGCTGCTCATCGAGATGATATATGCG	540
482	QY		
843	Db	ATGATGTACATCTCTCTGCTCTTCTCACCCTGTGGCTGCTGTTTATTGAGATGATCTATTGC	902
844	Db		
541	QY	TACAGAAAGTCTCAAAGCCGAGAGGCGAGCCCAAGAAAACCGCTCTGACTTACCTTCCG	600
542	QY		
903	Db	TACAGAAAGTCTCTTAGGCCGAGAGGCCAGCAGAGAAAATGGCTGACTTACCTTGCT	962
904	Db		
601	QY	ATCCCATCTGAGAACCAAGAGAACTCTTGGGTACCAAGTGGAGGAATA	647
602	QY		
963	Db	ATCCCTCTCAGAGAACCAAGGAACTCTTGGTACTCTGTGGAGGAATA	1009
964	Db		

```

RESULT 3
US-10-029-191-22
; Sequence 22, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, ROY A.J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 210147.00XX/SUI
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 2632
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-22

```

Query Match	82.8%;	Score 536.6;	DB 13;	Length 2632;
Best Local Similarity	89.3%;	Pred. No. 1.7e-161;		
Matches 578;	Conservative 0;	Mismatches 69;	Indels 0;	Gaps 0;
QY	1	ATGCCTGCCCTTCAATAGATTGTTCCCGTGGCTTCTCTCGTCTTATCTACTGGGTCAGT	60	
Db	78	ATGCCTGCCCTTCAACAGATTGCTCCCTAGCTTCTTAGTGTCTATCTACTGGGTCAGA	137	
QY	61	GTCTGCTTCCCTGTGTGTGTAAGTCCCTCGGAGACGGAGCGCGTGCAGGGCAACCCC	120	
Db	138	GTCTGCTTCCCTGTGTGTGTAAGTGCCTCGGAGACAGAGCGGTGCAGGGCAATCCC	197	
QY	121	ATGAGCTGGCTGCATCTCCTGCATGAGAGAGAGAGGTGGAGGCCACCACTGGTGTG	180	
Db	198	ATGAGCTGAGTGTGATCTCCTGCATGAGAGAGAGAGGTGGAGGCCACCACTGTGGTGTG	257	



181	DB	GAGTGGTTCTACAGGCCCTGAGGGCGGTAAAGATTTCCTTATATATGAGTATCGGAATGGC	240
241	QY	CACCAAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTCGAGTGGGAATGGCAGCAAGACCTG	300
241	DB		300
241	DB	CACCAAGGAGTGGAGAGCCCTTCCAAAGGCCGCTCGCAGTGGGAATGGGAGCAAAAGACCTG	300
301	QY	CAGGACGTGTCATCACTGTGTCTAAACGTCACTCTGAAAGACTCTGTGCCTCTACACCTGC	360
301	DB		360
361	QY	AATGTCTCCCGGAGTTTGTAGTTGTAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG	420
361	DB		420
421	QY	ATCCCCCTAAGATCACCGAGGAGCTCGAGAGCACTTCACTCTGTGGTCTCAGAAATC	480
421	DB		480
481	QY	ATGATGTACATCCTTCTGGTCTTCTCAACCTGTGGCTGCTCATCGAGATGATATTGCG	540
481	DB		540
541	QY	TACAGAAAGGTCTCAAAAGCCGGAAGAGCAGCCCAAGAAAACGCGTCTGACTACCTTGGC	600
541	DB		600
601	QY	ATCCCATCTGAGAACAGAGAGACTCTCGCGTTACAGTGGAGGAA	645
601	DB		645

## RESULT. T 6

```

RESULTS 6
US-10-029-191-23
/ Sequence 23, Application US/10029191
/ Publication No. US20020160453A1
/ GENERAL INFORMATION:
/ APPLICANT: CURTIS, ROY A.J.
/ TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
/ TITLE OF INVENTION: PROTEIN
/ FILE REFERENCE: 210147.00XX/5UI
/ CURRENT APPLICATION NUMBER: US/10/029,191
/ CURRENT FILING DATE: 2001-12-20
/ PRIOR APPLICATION NUMBER: 09/569,978
/ PRIOR FILING DATE: 2000-05-12
/ PRIOR APPLICATION NUMBER: US 60/134,198
/ PRIOR FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 23
/ LENGTH: 657
/ TYPE: DNA
/ ORGANISM: Lepus Sp.
US-10-029-191-23

```

QY	298	CTGCGAGACGCTGCCATCACTGTGCTCTCAAGCTACTCTGAACACACTCTGCCTCTACACC	355
DB	301	CTGCGAGACCTGTGCATCTTCATCACCAAATGTCACTTCAACACACTCGGGCGMCTACAG	360
QY	358	TGCAATGTGTCCCGGGAGCTTTGAGTCTTTGAGCGGCATCGGCCCTTTTGTGAGACGACGCGG	417
DB	361	TGCCATGTCTACCGCTCTCTCTCGAAAACTACGAGCAACAACACCGCTGCTCAAG	420
QY	418	CTGATCCCCCTAAGATGTCACGAGGAGCTGGAGAGACTTCACCTCTGTGGTCTCAGAA	477
DB	421	AAGATCCACCTGGAGGTGGTGGCAAGAGGCCAAOAGAGACATGCATCCATTCGTGTCCGAG	480
QY	478	ATCATGATGTATCATCTTCTTGCTTCTCTCAACCTGTGGCTGTCTATCGAGATGATATAT	537
DB	481	ATCATGATGTACGTCTCATCGTGGTGTGACCATCTGGCTCTGTGGCGGAGATGGTGTAC	540
QY	538	TGCTACAGAAAGGTCTCAAAAGCCGAGAGSCA---GCCCAAGAAACCGCTCTGACTAC	594
DB	541	TGCTACAGAAAGATTCGGCGGCCGACGAGAGGAGCGGGCGAGGAGAACGCTTCGGATATAC	600
QY	595	CTTGCCATCCCATCTTGAGAACCAAGAGAACTCTTGCGG	631
DB	601	TTGGCCATCACTCTCAGAAAGCAAGAAATTCACGG	637

## RESULT 7

```

US-09-917-800A-1654
; Sequence 1654, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Blashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44321-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1654
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_017388
US-09-917-800A-1654

```

76 TGTGTGAAGTGCCTCGAGAGCGAGCGCGTGCAGGCAACCCCATGAAGCTGGCTGC 135  
Db |||||  
280 TGGGTGAGGTGGATTCTGAGACCGAGGCGAGTGTATGGATGACCTTCAAAATCTGTGT 339  
Qy |||||  
136 ATCTCTCCATGAGAGAGAGGTGGAGGCCACACCGTGGTGAATGTTCTACAGG 195  
Db |||||  
340 ATCTCTCTGAAGCTGCTAGTGAGACACCGCGGAGACCTTACAGGAGTGGACCTTCGCG 399  
Qy |||||  
196 CCCGAGGCGCGTAAGATT-----TCTTTATTATCGAGTATCGGAATGCCACACG 246  
Db |||||  
400 CAGAGGCGCAGAGGAATTTGCAAGATCTTACGCTATGAGATGAGGTGCTGCGCTG 459  
Qy |||||  
247 GAGTGGAGAGCCCTTTGAGGGCGCTGCTGAGTGAATGGGAG-----CAAGAC 297  
Db |||||  
460 GAGGAAGATGAGCCCTTTGAGGGCGCTGCTGAGTGAACGTTAGTGGGGCGACCAAGGAC 519  
Qy |||||  
298 CTGAGGAGCGTGTCCATCACTGTGCTCAAGCTCACTCTGAACGACCTCTGCTCTACACC 357  
Db |||||  
520 CTGAGGAGCGTGTCCATCTTCACTACCAATGTCACTACCACTCTGGGACTAGGAA 579  
Qy |||||  
358 TGCATGTGTCGGGGAGTTGAGTTGAGGCGCATCGGCCCTTTGTGAGAGCGCGG 417  
Db |||||  
580 TGTCACTGTCACCT 639  
Qy |||||  
418 CTGATCCCTTAAAGATCAACGAGGAGGCTGGAGAGGACTTCACTCTCTCTCTCTCTCT 477  
Db |||||  
640 AAGATCCACCTGGAGGTGGTGGACAGGCGCAACAGAGATATGATCATCTCTCTCTCTCT 699  
Qy |||||  
478 ATCATGATGATCATCT 537  
Db |||||  
700 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759  
Qy |||||  
538 TGCTACAGAAAGGTCTCAAAAGCC--GAAGAGGCGAGCCCAAGAAACGCTCTGACTAC 594  
Db |||||  
760 TGCTACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 819  
Qy |||||  
595 CTGCTCATTCATCTGAGAACAGAGGAACTCTGCGG 631  
Db |||||  
820 CTGCGCATTAATCTCCGAGCGCAAGAGACTGTACAG 856

## RESULT a

US-10-029-386-16214  
; Sequence 16214, Application US/10029386  
; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: AEMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

; SEQ ID NO 16214

; LENGTH: 243

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR19.1

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2

; OTHER INFORMATION: SWISSPROT HIT: Q07699, EVALUE 5.00e-42

; OTHER INFORMATION: NT HIT: U12191.1, EVALUE 0.00e+00

; OTHER INFORMATION: EST\_HUMAN HIT: BF953777.1, EVALUE 1.00e-129

US-10-029-386-16214

## Query Match

Best Local Similarity 7.6%; Score 49; DB 12; Length 243;

Matches 104; Conservative 0; Mismatches 60; Indels 9; Gaps 1;

Qy 223 TAGAGTATCGGAATGGCCACGAGGAGTGGAGAGCCCTTTTCAGGGGCGCGCTGCAGTGG 282

Db |||||

4 TATGAGATGAGTGTTCGAGCTGGAGAGGATGAGCGCTTCGAGGGCGCGCTGTGTGG 63

Qy |||||

283 AATGGGAG-----CAAGGACCTGCGAGGAGTGTCCATCACTGTGTCAACGCTCACT 333

Db |||||

64 AATGGGAGCGGGGCGGACCAAGACCTGCGAGGATCTGTCTATCTCATCAACCAATGTCA 123

Qy |||||

334 CTGAACGACTCTGGCTCTACACCTCAATGTGTCCGGGAGTTGAGTTTGA 386

Db |||||

124 TACAACCACTCGGGGCGAGTACGAGTCCACGCTCTACCGGCTCTCTTCTTGA 176

US-10-029-386-2514

; Sequence 2514, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: AEMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

; SEQ ID NO 2514

; LENGTH: 569

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR19.1

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2

; OTHER INFORMATION: SWISSPROT HIT: Q07699, EVALUE 6.00e-42

; OTHER INFORMATION: NT HIT: U12191.1, EVALUE 0.00e+00

; OTHER INFORMATION: EST\_HUMAN HIT: BF754689.1, EVALUE 0.00e+00

US-10-029-386-2514

## Query Match

Best Local Similarity 7.6%; Score 49; DB 12; Length 569;

Matches 104; Conservative 0; Mismatches 60; Indels 9; Gaps 1;

Qy 223 TAGAGTATCGGAATGGCCACGAGGAGTGGAGAGCCCTTTTCAGGGGCGCGCTGCAGTGG 282

Db |||||

62 TATGAGATGAGTGTTCGAGCTGGAGAGGATGAGCGCTTCGAGGGCGCGCTGTGTGG 121

Qy |||||

283 AATGGGAG-----CAAGGACCTGCGAGGAGTGTCCATCACTGTGTCAACGCTCACT 333

Db |||||

122 AATGGGAGCGGGGCGGACCAAGACCTGCGAGGATCTGTCTATCTCATCAACCAATGTCA 181

Qy |||||

334 CTGAACGACTCTGGCTCTACACCTCAATGTGTCCGGGAGTTGAGTTTGA 386

Db |||||

182 TACAACCACTCGGGGCGAGTACGAGTCCACGCTCTACCGGCTCTCTTCTTGA 234

US-10-029-386-2514

## RESULT 10

US-10-094-466-63

; Sequence 63, Application US/10094466

; Publication No. US20030203363A1

; GENERAL INFORMATION:







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:16:11 ; Search time 1565.42 Seconds  
(without alignments) 9799.969 Million cell updates/sec

Title: US-09-936-680-4 COPY 1 375

```

Perfect score: 375
Sequence: 1 ccctccattccgagctgagc.....gaagccgcagcccccaaaq 375

```

Scoring table: IDENTITY\_NUC  
Gapox 10.0 Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_pa:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pri:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_pa:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pac:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_fun:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*\*

33: em\_htg\_mus:\*\*

34: em\_htg\_pln:\*\*

35: em\_htg\_rod:\*\*

36: em\_htg\_man:\*\*

37: em\_htg\_vrt:\*\*

38: em\_sy:\*\*

39: em\_htg\_hum:\*\*

40: em\_htg\_mus:\*\*

41: em\_htg\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	375	100.0	1261	6	AX39100	Sequence
2	375	100.0	1261	9	HSAX3396	AX243396 Homo sapi
3	373	99.5	127347	2	AP000682	AP000682 Homo sapi
4	373	99.5	144833	2	AC063921	AC063921 Homo sapi
5	373	99.5	143800	2	AC021981	AC021981 Homo sapi
6	373	99.5	172546	2	AC024604	AC024604 Homo sapi
7	373	99.5	178169	9	AP002765	AP002765 Homo sapi
8	373	99.5	181471	9	AC069539	AC069539 Homo sapi
9	369.8	98.6	4052	9	HSM801563	AL136589 Homo sapi
10	350	93.3	178431	2	AP002749	AP002749 Homo sapi
11	122	32.5	66980	2	AC013796	AC013796 Homo sapi
12	74	19.7	66980	2	AC013796	AC013796 Homo sapi
13	53	14.1	521	11	PM12B8G	AL684326 Penicill
14	48.2	12.9	200738	2	BX324126	BX324126 Danio rer
15	46.4	12.4	148643	2	AC011871	AC011871 Homo sapi
16	45.6	12.2	216449	2	AC051616	AC051616 Mus muscu
17	45.4	12.1	143624	2	AC135353	AC135353 Mus muscu
18	45.4	12.1	209257	2	AC123927	AC123927 Mus muscu
19	45.4	12.1	239130	2	AC079420	AC079420 Mus muscu
20	44.8	11.9	11608	9	HSINT62	X14445 Human int-2
21	44.8	11.9	149051	2	AP006345	AP006345 Homo sapi
22	44.6	11.9	58293	8	AC096781	AC096781 Oryza sat
23	44.6	11.9	117157	8	AC034258	AC034258 Oryza sat
24	44.6	11.9	197921	2	BX004847	BX004847 Danio rer
25	44.6	11.9	306161	8	AE017102	AE017102 Oryza sat
26	44.4	11.8	59364	2	AC102347	AC102347 Mus muscu
27	44.2	11.8	3252	9	HSTAP113	Y11354 H.sapiens m
28	44	11.7	85434	2	AC066610	AC066610 Homo sapi
29	43.8	11.7	152347	2	AC021096	AC021096 Homo sapi
30	43.6	11.6	972	11	PM12A13G	AL684288 Penicilli
31	43.6	11.6	1094	11	PM7G11B	AL685196 Penicilli
32	43.6	11.6	53897	2	AC101434	AC101434 Mus muscu
33	43.6	11.6	191983	9	AX744858	AX744858 Homo sapi
34	43.4	11.6	33676	1	SCARD1GN	X84374 Saccharothr
35	43.4	11.6	195039	2	AC013385	AC013385 Homo sapi
36	43.4	11.6	150118	2	AC137151	AC137151 Mus muscu
37	43.4	11.6	192888	2	BX005296	BX005296 Danio rer
38	43.4	11.6	240411	2	AC105428	AC105428 Mus muscu
39	43.2	11.5	61967	2	AC100100	AC100100 Mus muscu
40	43.2	11.5	64970	2	AC099958	AC099958 Mus muscu
41	43.2	11.5	204478	2	AC107063	AC107063 Bos tauru
42	43.2	11.5	239130	2	AC079420	AC079420 Mus muscu
43	43	11.5	87283	2	AC022856	AC022856 Homo sapi
44	42.8	11.4	110000	2	AC021632	AC021632 Mus muscu
45	42.6	11.4	59237	2	AC091279	AC091279 Mus muscu

## ALIGNMENTS

RESULT 1					
AX039100					
LOCUS					
DEFINITION					
AX039100					
ACCESSION					
AX039100.1					
VERSION					
GI:11229276					
KEYWORDS					
SOURCE					
ORGANISM					
Homo sapiens					
Homo sapiens					
Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
1					
Cox, P.; Dixon, A.; Jackson, A. and Morgan, K.					
A novel family of beta sub-unit proteins from a voltage-gated sodium channel, nucleic acids encoding them and therapeutic or					
REFERENCE					
AUTHORS					
TITLE					

diagnostic uses there of  
Patent: WO 0043367-A 4 26-OCT-2000;  
WARNER-LAMBERT COMPANY (US) ; Cambridge University Technical  
Services Limited (GB)  
FEATURES  
Location/Qualifiers  
source  
1..1261  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 263 a 366 c 372 g 260 t  
ORIGIN  
Query Match 100.0%; Score 375; DB 6; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 9.8e-70;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTCTCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGCGCGAGTGG 60  
Db 1 CCTCTCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGCGCGAGTGG 60  
QY 61 AAGCTGGAGTTCCTGGGCGGCGGAGCGGAGCTGCTCGGTGCTGAGCGCGCGGAG 120  
Db 61 AAGCTGGAGTTCCTGGGCGGCGGAGCGGAGCTGCTCGGTGCTGAGCGCGCGGAG 120  
QY 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 180  
Db 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 180  
QY 181 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 240  
Db 181 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 240  
QY 241 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 300  
Db 241 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 300  
QY 301 AAGAAATCTGAGAGCGGCGAGTCTTGACCGAGGGAATCTCTCTGTGTAGCTTTGGAAGC 360  
Db 301 AAGAAATCTGAGAGCGGCGAGTCTTGACCGAGGGAATCTCTCTGTGTAGCTTTGGAAGC 360  
QY 361 GCGGCGCGGAGCGGAG 375  
Db 361 GCGGCGCGGAGCGGAG 375  
RESULT 2  
HSA243396 1261 bp mRNA linear PRI 13-MAR-2000  
LOCUS  
DEFINITION Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit  
(scn3b gene).  
ACCESSION  
VERSION AJ243396.2 GI:7242612  
KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K.,  
Pinnock, R.D., Hughes, J., Richardson, P.J., Mizuguchi, K. and  
Jackson, A.P.  
beta 3: an additional auxiliary subunit of the voltage-sensitive  
sodium channel that modulates channel gating with distinct kinetics  
Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)  
20160948  
FUBMED 10688874  
REFERENCE  
2  
Morgan, K.  
Direct Submission  
Submitted (28-JUN-1999) Morgan K., Biochemistry, University of  
Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
Revised by [4]  
3 (bases 1 to 1261)

AUTHORS Morgan, K.  
Direct Submission  
TITLE Submitted (13-MAR-2000) Morgan K., Biochemistry, University of  
JOURNAL Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
COMMENT On Mar 14, 2000 this sequence version replaced GI:7160974.  
FEATURES Location/Qualifiers  
source  
1..1261  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue type="brain"  
1..1261  
/gene="scn3b"  
/db\_xref="scn3b"  
/gene="scn3b"  
376..1024  
/gene="scn3b"  
/function="modulates channel gating kinetics"  
/codon\_start=1  
/evidence=experimental  
/product="voltage-gated sodium channel beta-3 subunit"  
/protein\_id="CAB76825.1"  
/db\_xref="GI:7160975"  
/translation="MPAPNRLFLPLASLILYVSVCPVCPVEPSETEAVQGNPMKLR  
CISMKREVEATVVENFYRPGGKDFLIYRNGHQEVESPFQRLQWNGSKLDQ  
VSTLVNLTNDSGLYTCNVSREFEFHPPFKVTKRLILPLRVTEAGEDFTSVVSEI  
MMYLLVFLTLWLLIEMIYCYRKVSKAEAAQENASDYLAIPSENKENSAPVVEE"  
1025..>1261  
/gene="scn3b"  
BASE COUNT 263 a 366 c 372 g 260 t  
ORIGIN  
Query Match 100.0%; Score 375; DB 9; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 9.8e-70;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTCTCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGCGCGAGTGG 60  
Db 1 CCTCTCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGCGCGAGTGG 60  
QY 61 AAGCTGGAGTTCCTGGGCGGCGGAGCGGAGTCTCGGTGCTGAGCGCGCGGAG 120  
Db 61 AAGCTGGAGTTCCTGGGCGGCGGAGCGGAGTCTCGGTGCTGAGCGCGCGGAG 120  
QY 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 180  
Db 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 180  
QY 181 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 240  
Db 181 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 240  
QY 241 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 300  
Db 241 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCGAGTGGGCTGCTTAG 300  
QY 301 AAGAAATCTGAGAGCGGCGAGTCTTGACCGAGGGAATCTCTCTGTGTAGCTTTGGAAGC 360  
Db 301 AAGAAATCTGAGAGCGGCGAGTCTTGACCGAGGGAATCTCTCTGTGTAGCTTTGGAAGC 360  
QY 361 GCGGCGCGGAGCGGAG 375  
Db 361 GCGGCGCGGAGCGGAG 375  
RESULT 3  
AP000682/c 127347 bp DNA linear HTG 16-AUG-2000  
LOCUS  
DEFINITION Homo sapiens chromosome 11 clone CMB9-32A1 map 11q24, WORKING DRAFT  
SEQUENCE, 8 unordered pieces.  
ACCESSION AP000682  
VERSION AP000682.3 GI:9844967  
KEYWORDS HTGS\_PHASE1; HTGS\_DRAFT.





Db 50053 AGAATCTGAGAGGCGCGCTCTTACCGAGGAGTCTCTGTGTGACCTTGAGAGCGG 49994  
QY 363 CCAGCCCCCAGAG 375  
Db 49993 CCAGCCCCCAGAG 49981

RESULT 5  
AC021981/c  
LOCUS  
DEFINITION Homo sapiens chromosome 11 clone RP11-11C15 map 11, WORKING DRAFT  
SEQUENCE, 29 unordered pieces.  
AC021981  
VERSION 2, GI:9280747  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Becker, R., Bada, F.,  
Boguslavsky, L., Bouckgeater, B., Brown, A., Burkett, G., Castle, A.,  
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
DeRellano, K., Dewar, K., Domino, M., Doyle, M., Feneator, J.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Lander, J. C., Lechoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPheters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,  
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rochman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (23-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 19, 2000 this sequence version replaced GI:6731285.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIER  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L1335  
Center clone name: L1 C 15  
----- Summary Statistics  
Sequencing vector: M13, M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 135317 bases at least Q40  
Consensus quality: 141872 bases at least Q30  
Consensus quality: 14404 bases at least Q20  
Insert size: 121000; agarose-fp  
Quality coverage: 4.2 in Q20 bases; agarose-fp  
Quality coverage: 3.5 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 29 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1441: contig of 1441 bp in length  
\* 1442 1541: gap of 100 bp  
\* 1542 2883: contig of 1342 bp in length  
\* 2884 2984: gap of 100 bp  
\* 2984 4649: contig of 1666 bp in length  
\* 4650 4749: gap of 100 bp  
\* 4750 6112: contig of 1363 bp in length  
\* 6113 6212: gap of 100 bp  
\* 6213 7413: contig of 1201 bp in length  
\* 7414 7513: gap of 100 bp  
\* 7514 9051: contig of 1537 bp in length  
\* 9051 9151: gap of 100 bp  
\* 9151 10945: contig of 1794 bp in length  
\* 10945 11045: gap of 100 bp  
\* 11045 13203: contig of 2159 bp in length  
\* 13204 13303: gap of 100 bp  
\* 13304 16118: contig of 2815 bp in length  
\* 16119 16218: gap of 100 bp  
\* 16219 18573: contig of 2455 bp in length  
\* 18574 18774: gap of 100 bp  
\* 18774 21181: contig of 2407 bp in length  
\* 21181 21281: gap of 100 bp  
\* 21281 23200: contig of 1920 bp in length  
\* 23201 23300: gap of 100 bp  
\* 23301 26526: contig of 3226 bp in length  
\* 26527 26626: gap of 100 bp  
\* 26627 28655: contig of 3029 bp in length  
\* 28656 29755: gap of 100 bp  
\* 29756 34739: contig of 4983 bp in length  
\* 34739 34838: gap of 100 bp  
\* 34839 39394: contig of 4856 bp in length  
\* 39395 39494: gap of 100 bp  
\* 39495 43669: contig of 4175 bp in length  
\* 43670 43769: gap of 100 bp  
\* 43770 49948: contig of 6179 bp in length  
\* 49949 50048: gap of 100 bp  
\* 50049 55578: contig of 5530 bp in length  
\* 55579 55678: gap of 100 bp  
\* 55679 60959: contig of 5281 bp in length  
\* 60960 61059: gap of 100 bp  
\* 61060 66394: contig of 5335 bp in length  
\* 66395 72948: contig of 6453 bp in length  
\* 72949 73047: gap of 100 bp  
\* 73048 80874: contig of 7827 bp in length  
\* 80875 80974: gap of 100 bp  
\* 80975 87847: contig of 6873 bp in length  
\* 87848 87947: gap of 100 bp  
\* 87948 97279: contig of 9332 bp in length  
\* 97280 97379: gap of 100 bp  
\* 97380 106037: contig of 8658 bp in length  
\* 106038 106137: gap of 100 bp  
\* 106138 114206: contig of 8069 bp in length  
\* 114207 114306: gap of 100 bp  
\* 114307 129830: contig of 15524 bp in length  
\* 129831 129931: gap of 100 bp  
\* 129931 149800: contig of 19870 bp in length.  
FEATURES  
source  
1. 149800  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-11C15"  
/clone\_lib="RP11-11 Human Male BAC"  
misc\_feature 1. 1441  
/note="assembly\_fragment"  
1542. .2883  
misc\_feature

Query Match 99.5%; Score 373; DB 2; Length 149800;  
Best Local Similarity 100.0%; Pred. No. 1.2e-69;

```
* 2790 2889: gap of unknown length
* 2890 4562: contig of 1673 bp in length
* 4563 4662: gap of unknown length
* 4663 6344: contig of 1682 bp in length
* 6345 6444: gap of unknown length
* 6445 9221: contig of 2777 bp in length
* 9222 9321: gap of unknown length
* 9322 11997: contig of 2676 bp in length
* 11998 12097: gap of unknown length
* 12098 14867: contig of 2770 bp in length
* 14868 14967: gap of unknown length
* 14968 19319: contig of 4952 bp in length
* 19320 20019: gap of unknown length
* 20020 24125: contig of 4106 bp in length
* 24126 24225: gap of unknown length
* 24226 29444: contig of 5219 bp in length
* 29445 29544: gap of unknown length
* 29545 34568: contig of 5024 bp in length
* 34569 34668: gap of unknown length
* 34669 39414: contig of 4746 bp in length
* 39415 39514: gap of unknown length
* 39515 45885: contig of 6371 bp in length
* 45886 50986: contig of 5012 bp in length
* 50987 50997: gap of unknown length
* 50998 59143: contig of 8046 bp in length
* 59144 59243: gap of unknown length
* 59244 67779: contig of 8536 bp in length
* 67780 67879: gap of unknown length
* 67880 76429: contig of 8550 bp in length
* 76430 76529: gap of unknown length
* 76530 86704: contig of 10175 bp in length
* 86705 86804: gap of unknown length
* 86805 103147: contig of 16343 bp in length
* 103148 103247: gap of unknown length
* 103248 116343: contig of 13096 bp in length
* 116344 116443: gap of unknown length
* 116444 135318: contig of 18875 bp in length
* 135319 135418: gap of unknown length
* 135419 172546: contig of 37128 bp in length.
FEATURES
    source
    1..172546
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="Chromosome 10"
    /clone_lib="RPC1-11"
    /clone="RP11-331G19"
BASE COUNT 47530 a 37712 c 38612 g 46556 t 2136 others
ORIGIN
Query Match 99.5%; Score 373; DB 2; Length 172546;
Best Local Similarity 100.0%; Pred. No. 1.1e-69;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CTCCTTCGAGCTGAGCTTACCTCGGGCGCAACGAGCGAGGAGGGCCGAGTGGAA 62
DB 122377 CTCCTTCGAGCTGAGCTTACCTCGGGCGCAACGAGCGAGGAGGGCCGAGTGGAA 122436
QY 63 GCTGGAGTTCCGGGCTGGGCGGGAGGCGACTGTCCGTGGTCTGAGCGCCGGCGAGAGC 122
DB 122437 GCTGGAGTTCCGGGCTGGGCGGGAGGCGACTGTCCGTGGTCTGAGCGCCGGCGAGAGC 122496
QY 123 GGGCGGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCTAGGG 182
DB 122497 GGGCGGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCTAGGG 122556
QY 183 CCCAAAGCCCAACCCGGGCTCCAAAGAGTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCGCGC 242
DB 122557 CCCAAAGCCCAACCCGGGCTCCAAAGAGTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCGCGC 122616
QY 243 CCTTCCTTCGGTCAGAAAGTCCGCCCTCGGGGCGAGTTCGTCCTCCCAAGGGTTTCTCGAA 302
122617 CTTTCTTCGGTCAGAAAGTCCGCCCTCGGGGCGAGTTCGTCCTCCCAAGGGTTTCTCGAA 122676
QY 303 AGAATCTGAGAGGGCGGAGTCTTTGACCGGAGGAATCTCTCTGTAGCTTCTGGAAGCCG 362
122677 AGAATCTGAGAGGGCGGAGTCTTTGACCGGAGGAATCTCTCTGTAGCTTCTGGAAGCCG 122736
QY 363 CCAGCCCAAG 375
DB 122737 CCAGCCCAAG 122749
RESULT 7
AP002765/c 178169 bp DNA linear PRI 15-MAR-2003
LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-634B22,
DEFINITION complete sequence.
ACCESSION AP002765
VERSION AP002765.3 GI:16751488
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsr.riken.go.jp, URL:http://hgp.gs.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Nov 5, 2001 this sequence version replaced gi:12381934.
FEATURES
    Location/Qualifiers
    1..178169
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11q"
    /clone="RP11-634B22"
BASE COUNT 46062 a 39655 c 41140 g 51312 t
ORIGIN
Query Match 99.5%; Score 373; DB 9; Length 178169;
Best Local Similarity 100.0%; Pred. No. 1.1e-69;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CTCCTTCGAGCTGAGCTTACCTCGGGCGCAACGAGCGAGGAGGGCCGAGTGGAA 62
DB 107533 CTCCTTCGAGCTGAGCTTACCTCGGGCGCAACGAGCGAGGAGGGCCGAGTGGAA 107474
QY 63 GCTGGAGTTCCGGGCTGGGCGGGAGGCGACTGTCCGTGGTCTGAGCGCCGGCGAGAGC 122
DB 107473 GCTGGAGTTCCGGGCTGGGCGGGAGGCGACTGTCCGTGGTCTGAGCGCCGGCGAGAGC 107414
QY 123 GGGCGGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCTAGGG 182
DB 107413 GGGCGGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCTAGGG 107354
QY 183 CCCAAAGCCCAACCCGGGCTCCAAAGAGTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCGCGC 242
DB 107353 CCCAAAGCCCAACCCGGGCTCCAAAGAGTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCGCGC 107294
QY 243 CCTTCCTTCGGTCAGAAAGTCCGCCCTCGGGGCGAGTTCGTCCTCCCAAGGGTTTCTCGAA 302
107293 CCTTCCTTCGGTCAGAAAGTCCGCCCTCGGGGCGAGTTCGTCCTCCCAAGGGTTTCTCGAA 107234
```







Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
Direct Submission  
TITLE  
JOURNAL  
COMMENT  
Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 27, 2001 this sequence version replaced gi:7582586.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L3904  
Center clone name: 20\_H\_9

\* NOTE: This record contains 83 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 683: contig of 683 bp in length  
\* 684 783: gap of 100 bp  
\* 784 1488: contig of 705 bp in length  
\* 1489 1588: gap of 100 bp  
\* 1589 2296: contig of 708 bp in length  
\* 2297 2396: gap of 100 bp  
\* 2397 3101: contig of 705 bp in length  
\* 3102 3201: gap of 100 bp  
\* 3202 3316: contig of 715 bp in length  
\* 3317 4016: gap of 100 bp  
\* 4017 4737: contig of 721 bp in length  
\* 4738 4837: gap of 100 bp  
\* 4838 5557: contig of 720 bp in length  
\* 5558 5657: gap of 100 bp  
\* 5659 6372: contig of 715 bp in length  
\* 6373 6472: gap of 100 bp  
\* 6473 7155: contig of 683 bp in length  
\* 7156 7255: gap of 100 bp  
\* 7256 7957: contig of 702 bp in length  
\* 7958 8057: gap of 100 bp  
\* 8058 8773: contig of 716 bp in length  
\* 8774 9377: contig of 704 bp in length  
\* 9378 9677: gap of 100 bp  
\* 9678 10392: contig of 715 bp in length  
\* 10393 10492: gap of 100 bp  
\* 10493 11179: contig of 687 bp in length  
\* 11180 11279: gap of 100 bp  
\* 11280 11984: contig of 705 bp in length  
\* 11985 12084: gap of 100 bp  
\* 12085 12791: contig of 707 bp in length  
\* 12792 12891: gap of 100 bp  
\* 12892 13600: contig of 709 bp in length  
\* 13601 13700: gap of 100 bp  
\* 13701 14409: contig of 709 bp in length  
\* 14410 14509: gap of 100 bp  
\* 14510 15221: contig of 712 bp in length  
\* 15222 15321: gap of 100 bp  
\* 15322 16025: contig of 704 bp in length  
\* 16026 16125: gap of 100 bp  
\* 16126 16822: contig of 697 bp in length  
\* 16823 16922: gap of 100 bp  
\* 16923 17621: contig of 699 bp in length  
\* 17622 17721: gap of 100 bp  
\* 17722 18423: contig of 702 bp in length  
\* 18424 18523: gap of 100 bp  
\* 18524 19241: contig of 718 bp in length  
\* 19242 19341: gap of 100 bp  
\* 19342 20056: contig of 715 bp in length  
\* 20057 20156: gap of 100 bp  
\* 20157 20870: contig of 714 bp in length  
\* 20871 20970: gap of 100 bp  
\* 20971 21671: contig of 701 bp in length  
\* 21672 21771: gap of 100 bp  
\* 21772 22480: contig of 709 bp in length  
\* 22481 22580: gap of 100 bp  
\* 22581 23293: contig of 713 bp in length  
\* 23294 23393: gap of 100 bp  
\* 23394 24095: contig of 702 bp in length  
\* 24096 24195: gap of 100 bp  
\* 24196 24916: contig of 721 bp in length  
\* 24917 25016: gap of 100 bp  
\* 25017 25727: contig of 710 bp in length  
\* 25728 25826: gap of 100 bp  
\* 25827 26534: contig of 708 bp in length  
\* 26535 26634: gap of 100 bp  
\* 26635 27333: contig of 699 bp in length  
\* 27334 27433: gap of 100 bp  
\* 27434 28131: contig of 698 bp in length  
\* 28132 28231: gap of 100 bp  
\* 28232 28926: contig of 695 bp in length  
\* 28927 29026: gap of 100 bp  
\* 29027 29758: contig of 732 bp in length  
\* 29759 29858: gap of 100 bp  
\* 29859 30566: contig of 708 bp in length  
\* 30567 30666: gap of 100 bp  
\* 30667 31368: contig of 702 bp in length  
\* 31369 31468: gap of 100 bp  
\* 31469 32183: contig of 715 bp in length  
\* 32184 32283: gap of 100 bp  
\* 32284 32981: contig of 698 bp in length  
\* 32982 33081: gap of 100 bp  
\* 33082 33798: contig of 717 bp in length  
\* 33799 33898: gap of 100 bp  
\* 33899 34634: contig of 736 bp in length  
\* 34635 34734: gap of 100 bp  
\* 34735 35433: contig of 699 bp in length  
\* 35434 35533: gap of 100 bp  
\* 35534 36253: contig of 720 bp in length  
\* 36254 36353: gap of 100 bp  
\* 36354 37079: contig of 726 bp in length  
\* 37080 37179: gap of 100 bp  
\* 37180 37893: contig of 714 bp in length  
\* 37894 37993: gap of 100 bp  
\* 37994 38700: contig of 707 bp in length  
\* 38701 38800: gap of 100 bp  
\* 38801 39505: contig of 705 bp in length  
\* 39506 39605: gap of 100 bp  
\* 39606 40316: contig of 711 bp in length  
\* 40317 40416: gap of 100 bp  
\* 40418 41131: contig of 715 bp in length  
\* 41132 41231: gap of 100 bp  
\* 41232 41339: contig of 708 bp in length  
\* 41340 42039: gap of 100 bp  
\* 42040 42748: contig of 709 bp in length  
\* 42749 42848: gap of 100 bp  
\* 42849 43555: contig of 707 bp in length  
\* 43556 43655: gap of 100 bp  
\* 43656 44368: contig of 713 bp in length  
\* 44369 44468: gap of 100 bp  
\* 44469 45171: contig of 703 bp in length  
\* 45172 45271: gap of 100 bp  
\* 45272 45943: contig of 672 bp in length  
\* 45944 46043: gap of 100 bp  
\* 46044 46749: contig of 706 bp in length  
\* 46750 46849: gap of 100 bp  
\* 46850 47550: contig of 701 bp in length  
\* 47551 47650: gap of 100 bp

```

* 47651 48350: contig of 700 bp in length
* 48351 48450: gap of 100 bp
* 48451 49148: contig of 698 bp in length
* 49149 49248: gap of 100 bp
* 49249 49961: contig of 713 bp in length
* 49962 50061: gap of 100 bp
* 50062 50751: contig of 690 bp in length
* 50752 50851: gap of 100 bp
* 50852 51567: contig of 716 bp in length
* 51568 52400: contig of 733 bp in length
* 52401 53225: contig of 725 bp in length
* 53226 53974: gap of 100 bp
* 53975 54074: contig of 649 bp in length
* 54075 54781: gap of 100 bp
* 54782 54881: contig of 707 bp in length
* 54882 55708: gap of 100 bp
* 55709 56426: contig of 718 bp in length
* 56427 57244: gap of 100 bp
* 57245 57344: contig of 718 bp in length
* 57345 58051: gap of 100 bp
* 58052 58151: contig of 707 bp in length
* 58152 58860: gap of 100 bp
* 58860: contig of 709 bp in length

Query Match 32.5%; Score 122; DB 2; Length 66980;
Best Local Similarity 99.3%; Pred. No. 2.7e-16;
Matches 133; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 242 CCTTCTTCGTGAGAGTGCCTCCCTGGGCGAGTTCGCCAAGGGTTCTCGA 301
DB 63886 CCTTCTTCGTGAGAGTGCCTCCCTGGGCGAGTTCGCCAAGGGTTCTCGA 63944

QY 302 AAGATCTGAGGGCGGAGTCCCTTACCGAGGGAATCTCTGTAGCCTTGGAGCC 361
DB 63945 AAGATCTGAGGGCGGAGTCCCTTACCGAGGGAATCTCTGTAGCCTTGGAGCC 64004

QY 362 GCCAGCCCGAAG 375
DB 64005 GCCAGCCCGAAG 64018

```

```

RESULT 12
AC013796/c
LOCUS
DEFINITION
AC013796
AC013796.4 GI:12583842
VERSION
HTG; HTGS PHASE0.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birken, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone RP11-20H9
Unpublished
2 (bases 1 to 66980)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Becker, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Collings, S., Collins, S., Collamore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McSwan, P., McGurk, A., McKernan, K., McLoughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

```

Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 27, 2001 this sequence version replaced gi:7582586.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L3904

Center clone name: 20\_H\_9

NOTE: This record contains 83 individual  
sequencing reads that have not been assembled into  
contigs. Runs of N are used to separate the reads  
and the order in which they appear is completely  
arbitrary. Low-pass sequence sampling is useful for  
identifying clones that may be gene-rich and allows  
overlap relationships among clones to be deduced.  
However, it should not be assumed that this clone  
will be sequenced to completion. In the event that  
the record is updated, the accession number will  
be preserved.

```

* 1 683: contig of 693 bp in length
* 684 783: gap of 100 bp
* 784 1488: contig of 705 bp in length
* 1489 1588: gap of 100 bp
* 1589 2296: contig of 708 bp in length
* 2297 3101: gap of 100 bp
* 3102 3201: contig of 705 bp in length
* 3202 3916: contig of 715 bp in length
* 3917 4016: gap of 100 bp
* 4017 4737: contig of 721 bp in length
* 4738 4837: gap of 100 bp
* 4838 5557: contig of 720 bp in length
* 5558 6372: gap of 100 bp
* 6373 6472: gap of 100 bp
* 6473 7155: contig of 683 bp in length
* 7156 7255: gap of 100 bp
* 7256 7957: contig of 702 bp in length
* 7958 8057: gap of 100 bp
* 8058 8774: contig of 716 bp in length
* 8774 9577: gap of 100 bp
* 9578 10392: contig of 704 bp in length
* 10393 10493: contig of 715 bp in length
* 10494 11179: gap of 100 bp
* 11180 11279: contig of 687 bp in length
* 11280 11984: contig of 705 bp in length
* 11985 12084: gap of 100 bp
* 12085 12791: contig of 707 bp in length
* 12792 12891: gap of 100 bp
* 12892 13600: contig of 709 bp in length
* 13601 13700: gap of 100 bp
* 13701 14409: contig of 709 bp in length
* 14410 14509: gap of 100 bp
* 14510 15221: contig of 712 bp in length
* 15222 15321: gap of 100 bp
* 15322 16025: contig of 704 bp in length
* 16026 16125: gap of 100 bp
* 16126 16822: contig of 697 bp in length
* 16823 17621: contig of 699 bp in length
* 17622 17721: gap of 100 bp

```

\* 17722 18423: contig of 702 bp in length  
\* 18424 18523: gap of 100 bp  
\* 18524 19241: contig of 718 bp in length  
\* 19241 19242: gap of 100 bp  
\* 19242 20056: contig of 715 bp in length  
\* 20056 20156: gap of 100 bp  
\* 20156 20870: contig of 714 bp in length  
\* 20870 20970: gap of 100 bp  
\* 20970 21671: contig of 701 bp in length  
\* 21671 21771: gap of 100 bp  
\* 21771 22480: contig of 709 bp in length  
\* 22480 22580: gap of 100 bp  
\* 22580 23293: contig of 713 bp in length  
\* 23293 23394: gap of 100 bp  
\* 23394 24095: contig of 702 bp in length  
\* 24095 24195: gap of 100 bp  
\* 24195 24916: contig of 721 bp in length  
\* 24916 25017: gap of 100 bp  
\* 25017 25727: contig of 710 bp in length  
\* 25727 26534: gap of 100 bp  
\* 26534 26635: contig of 708 bp in length  
\* 26635 27333: contig of 693 bp in length  
\* 27333 27434: gap of 100 bp  
\* 27434 28131: contig of 698 bp in length  
\* 28131 28232: gap of 100 bp  
\* 28232 28926: contig of 695 bp in length  
\* 28926 29027: gap of 100 bp  
\* 29027 29758: contig of 732 bp in length  
\* 29758 30566: gap of 100 bp  
\* 30566 31368: contig of 702 bp in length  
\* 31368 31469: gap of 100 bp  
\* 31469 32183: contig of 715 bp in length  
\* 32183 32884: gap of 100 bp  
\* 32884 32981: contig of 698 bp in length  
\* 32981 33082: gap of 100 bp  
\* 33082 33798: contig of 717 bp in length  
\* 33798 33899: gap of 100 bp  
\* 33899 34635: contig of 736 bp in length  
\* 34635 34734: gap of 100 bp  
\* 34734 35333: contig of 639 bp in length  
\* 35333 36253: gap of 100 bp  
\* 36253 36354: contig of 720 bp in length  
\* 36354 37079: contig of 726 bp in length  
\* 37079 37179: gap of 100 bp  
\* 37179 37893: contig of 714 bp in length  
\* 37893 37994: gap of 100 bp  
\* 37994 38700: contig of 707 bp in length  
\* 38700 38800: gap of 100 bp  
\* 38800 39505: contig of 705 bp in length  
\* 39505 39606: gap of 100 bp  
\* 39606 40317: contig of 711 bp in length  
\* 40317 41131: gap of 100 bp  
\* 41131 41231: contig of 715 bp in length  
\* 41231 41939: contig of 708 bp in length  
\* 41939 42039: gap of 100 bp  
\* 42039 42748: contig of 709 bp in length  
\* 42748 42849: gap of 100 bp  
\* 42849 43555: contig of 707 bp in length  
\* 43555 43656: gap of 100 bp  
\* 43656 44368: contig of 713 bp in length  
\* 44368 44469: gap of 100 bp  
\* 44469 45171: contig of 703 bp in length  
\* 45171 45272: gap of 100 bp  
\* 45272 45944: contig of 672 bp in length  
\* 45944 46043: gap of 100 bp  
\* 46043 46749: contig of 706 bp in length  
\* 46749 46750: gap of 100 bp  
\* 46750 46850: contig of 701 bp in length

\* 47551 47650: gap of 100 bp  
\* 47651 48350: contig of 700 bp in length  
\* 48351 48450: gap of 100 bp  
\* 48451 49148: contig of 698 bp in length  
\* 49149 49248: gap of 100 bp  
\* 49249 49961: contig of 713 bp in length  
\* 49962 50061: gap of 100 bp  
\* 50062 50751: contig of 690 bp in length  
\* 50752 50851: gap of 100 bp  
\* 50852 51567: contig of 716 bp in length  
\* 51568 51667: gap of 100 bp  
\* 51668 52400: contig of 733 bp in length  
\* 52401 53225: contig of 725 bp in length  
\* 53226 53325: gap of 100 bp  
\* 53326 53974: contig of 649 bp in length  
\* 53975 54074: gap of 100 bp  
\* 54075 54781: contig of 707 bp in length  
\* 54782 54881: gap of 100 bp  
\* 54882 55708: contig of 727 bp in length  
\* 55609 55708: gap of 100 bp  
\* 55709 56427: contig of 718 bp in length  
\* 56427 56527: gap of 100 bp  
\* 56527 57244: contig of 718 bp in length  
\* 57245 57344: gap of 100 bp  
\* 57345 58051: contig of 707 bp in length  
\* 58052 58151: gap of 100 bp  
\* 58152 58860: contig of 709 bp in length

Query Match 19.7%; Score 74; DB 2; Length 66980;  
Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 302 AAGAACTGAGGCGCGAGTCTTGCACCGAGGGAATCTCTGTGTAGCTTGGAGCC 361  
|||||  
Db 48350 AAGAACTGAGGCGCGAGTCTTGCACCGAGGGAATCTCTGTGTAGCTTGGAGCC 48291  
|||||

Qy 362 GCCAGCCCCAGAG 375  
|||||  
Db 48290 GCCAGCCCCAGAG 48277  
|||||

RESULT 13  
PM12B8G/c  
LOCUS PM12B8G 521 bp DNA linear STS 09-MAR-2002  
DEFINITION Penicillium marneffe STS, clone pm12b8.g, sequence tagged site.  
ACCESSION AL684326  
VERSION AL684326.1 GI:193337430  
KEYWORDS STS.  
SOURCE Penicillium marneffe  
ORGANISM Penicillium marneffe  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.  
REFERENCE 1 Yuen,K.Y., Pascal,G., Wong,S., Glaser,P., Woo,P., Kunst,P.,  
AUTHORS Cheung,E., Medigue,C. and Danchin,A.  
TITLE Exploring the Penicillium marneffe genome  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 521)  
AUTHORS Danchin,A. and Pascal,G.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAR-2002) Danchin A., HKU-Pasteur Research Centre,  
Dexter HC Man Building 8, Sassoon Road, Pokfulam, Hong Kong

FEATURES  
source  
1..521  
/organism="Penicillium marneffe"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3772"  
/clone="pm12b8.g"

BASE COUNT 29 a 287 c 192 g 9 t 4 others  
ORIGIN

Query Match 14.1%; Score 53; DB 11; Length 521;

Best Local Similarity 51.0%; Pred. No. 0.27;  
 Matches 122; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 48 GGGGCGCGAGTGAAGCTGGAGTTCGGGTGGGGGAGGAGCGACTGTCGTGGTGTG 107  
 |||  
 Db 317 GGGGCGCGAGTGAAGCTGGAGTTCGGGTGGGGGAGGAGCGACTGTCGTGGTGTG 258  
 |||

QY 108 AGCGCGCGAGAGCGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 167  
 |||  
 Db 257 GCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 198  
 |||

QY 168 TGGGGTGGCTTAAAGCGCCAAAGCCCGCCAGCTCCAAAGCTCCAGGGCTCCCGCAG 227  
 |||

Db 197 GGG 138  
 |||

QY 228 GCACCGGTCTGCGCCCTTCCTTGGTTCAGAAAGTGGCCCTGGGGGGGAGTTCGTCC 286  
 |||

Db 137 GCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 79  
 |||

RESULT 14  
 BX324126/c  
 LOCUS  
 DEFINITION  
 accession  
 version  
 keywords  
 source  
 organism

BX324126 200738 bp DNA linear HTG 03-MAY-2003  
 Danio rerio clone CH211-46H12, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 8  
 unordered pieces.

BX324126  
 HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 Danio rerio (zebrafish)  
 Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 200738)

McLay, K.  
 Direct Submission  
 Submitted (02-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On May 3, 2003 this sequence version replaced gi:30089190.

----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: zfish-help@sanger.ac.uk  
 ----- Project Information  
 Center project name: z46H12  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Chemistry: Dye-terminator; 100% of reads  
 Consensus quality: 198105 bases at least Q40  
 Consensus quality: 198375 bases at least Q30  
 Consensus quality: 198535 bases at least Q20  
 Insert size: 200038; sum-of-contigs  
 Insert size: 177373; 7.2% error; agarose-fp  
 Quality coverage: 7.51x in Q20 bases; sum-of-contigs Quality  
 coverage: 8.47x in Q20 bases; agarose-fp

## COMMENT

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 35213: contig of 35213 bp in length  
 \* 35214 35313: gap of 100 bp  
 \* 35314 74575: contig of 39262 bp in length  
 \* 74576 74676: gap of 100 bp  
 \* 74676 87157: contig of 12482 bp in length  
 \* 87158 87257: gap of 100 bp  
 \* 87258 128594: contig of 41337 bp in length

\* 128595 128694: gap of 100 bp  
 \* 128695 14799: contig of 13285 bp in length  
 \* 147980 148079: gap of 100 bp  
 \* 148080 151034: contig of 2955 bp in length  
 \* 151035 151134: gap of 100 bp  
 \* 151135 170296: contig of 19162 bp in length  
 \* 170297 170397: gap of 100 bp  
 \* 170397 200738: contig of 30342 bp in length.

## FEATURES

source  
 1..200738  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="CH211-46H12"  
 /clone\_lib="CHORI-211"  
 1..35213  
 /note="assembly fragment:02575"  
 fragment\_chain:1  
 clone\_end:SP6  
 vector\_side:left  
 35314..74575  
 /note="assembly fragment:01950"  
 fragment\_chain:1  
 74676..87157  
 /note="assembly fragment:00676"  
 fragment\_chain:1  
 87258..128594  
 /note="assembly fragment:02426"  
 fragment\_chain:1  
 128695..147979  
 /note="assembly fragment:01285"  
 fragment\_chain:1  
 148080..151034  
 /note="assembly fragment:01853"  
 fragment\_chain:1  
 151135..170296  
 /note="assembly fragment:00968"  
 fragment\_chain:1  
 170397..200738  
 /note="assembly fragment:02576"  
 clone\_end:T7  
 vector\_side:right

BASE COUNT 60312 a 40395 c 40220 g 59095 t 716 others  
 ORIGIN

Query Match 12.9%; Score 48.2; DB 2; Length 200738;  
 Best Local Similarity 60.2%; Pred. No. 1.1;  
 Matches 80; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 40 GCGAGGCGAGGCGGCGGAGTGGAGCTGAGTTCGGGGTGGGCGGAGGCGGAGTGTCCG 99  
 |||

Db 128473 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 128414  
 |||

QY 100 TGGTGTGAGCGCGCGGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 159  
 |||

Db 128413 CGGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 128354  
 |||

QY 160 AGTTCAGTGGGG 172  
 |||

Db 128353 CGGGCGCGCGCGG 128341  
 |||

## RESULT 15

AC011871/c  
 LOCUS  
 DEFINITION  
 accession  
 version  
 keywords  
 source  
 organism

AC011871 148643 bp DNA linear HTG 12-MAR-2000  
 Homo sapiens clone RP11-1618, WORKING DRAFT SEQUENCE, 14 unordered  
 pieces.

AC011871  
 HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





GenCore version 5.1.6  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:12:31; Search time 141.257 Seconds  
(without alignments)  
7166.298 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_1\_375

Perfect score: 375

Sequence: 1 cctcctctccagctgagc.....gaagccgagccagccagagaag 375

Scoring table: IDENTITY\_NUC

Gapop 10.0, Capext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

N\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375	100.0	1261	21	Human beta3 CDNA.
2	369.8	98.6	4052	24	Human signal trans
3	288.4	76.9	978	22	Human EST-derived
4	286.8	76.5	1045	22	Human polynucleoti
5	40.2	10.7	2806	20	Full length human
6	40.2	10.7	2811	23	DNA encoding novel
7	39.8	10.6	392	25	Bovine EST associa
8	39.6	10.6	2685	24	Human cDNA differe

9	39	10.4	936	24	ABK73545
10	39	10.4	3272	24	AB199475
c 11	38.6	10.3	1337	20	AA217263
12	38.6	10.3	1776	22	AA06422
13	38.6	10.3	1941	21	AA240657
c 14	38.6	10.3	2310	25	ABX70915
c 15	38.4	10.2	1544	21	AA076504
c 16	38	10.1	2307	24	ABK32804
17	37.8	10.1	802	24	AB161151
18	37.8	10.1	1275	11	AAQ01620
19	37.8	10.1	1344	12	AAQ13318
20	37.8	10.1	1345	11	AAQ03062
21	37.8	10.1	1495	22	AAF84455
22	37.8	10.1	4211	22	ABH21793
23	37.8	10.1	4286	24	ABN85772
24	37.8	10.1	4778	25	ACC44674
25	37.8	10.1	5504	24	AA143172
26	37.8	10.1	5510	25	ACC44675
27	37.8	10.1	5565	24	AA143171
28	37.8	10.1	5581	22	AAH41035
29	37.8	10.1	5630	25	AL50223
30	37.8	10.1	5759	24	ABK49521
31	37.8	10.1	5855	25	ACC44715
32	37.8	10.1	5855	25	ACC44730
33	37.8	10.1	5855	25	ABT16616
c 34	37.8	10.1	5878	24	ABT08199
35	37.8	10.1	6119	25	ACC44729
36	37.8	10.1	6148	24	ABK49520
37	37.8	10.1	6256	24	AA143173
38	37.8	10.1	6641	24	ABT08200
39	37.8	10.1	7600	25	ACC44718
40	37.8	10.1	7631	25	ACC44719
41	37.8	10.1	8031	22	AD12739
42	37.8	10.1	8521	25	ACC44726
43	37.8	10.1	8531	25	ACC44727
44	37.8	10.1	9080	25	ACC44713
45	37.8	10.1	10474	25	ACC44728

## ALIGNMENTS

### RESULT 1

AA067837 standard; cdna; 1261 BP.

XX AAC67837;

XX 15-FEB-2001 (first entry)

XX Human beta3 CDNA.

XX Human; beta sub-unit; beta3; analgesic; anticonvulsant;  
cerebroprotective; vasotropic; cardiast; nootropic; cytosstatic;  
dermatological; gene therapy; voltage-gated sodium channel; pain;  
epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;  
familial nonchromaffin paraganglioma; phenylketonuria;  
Charcot Marie tooth disease; ss.

XX Homo sapiens.

XX WO200063367-A1.

XX 26-OCT-2000.

XX 24-FEB-2000; 2000WO-EP01783.

XX 15-APR-1999; 99US-0129473.

XX (WARN ) WARNER LAMBEET CO.

XX (UTCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Cox P, Dixon A, Jackson A, Morgan K;

XX WPI: 2000-665241/64.  
 DR P-PSDB; AAB36002.  
 XX  
 XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
 PT sodium channel, and their corresponding polypeptides, useful for  
 PT detecting and treating sodium channel-associated conditions, e.g. pain,  
 PT epilepsy and stroke  
 XX  
 XX Claim 10; Page 70-71; 88pp; English.  
 PS  
 XX The present sequence is given in the claims of a specification  
 CC relating to a novel family of beta sub-unit proteins from a  
 CC voltage-gated sodium channel. Human and rat beta sub-units, which  
 CC have been collectively identified as beta3, have been isolated.  
 CC The polynucleotides and polypeptides are useful for screening for  
 CC agonists and antagonists of sodium channels. The agonists, antagonists,  
 CC proteins and nucleic acids may be used diagnosing of treating diseases,  
 CC or conditions associated with voltage-gated sodium channels, e.g. pain,  
 CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial  
 CC Nonchromaffin Paraganglioma, Phenylketonuria and Charcot Marie Tooth  
 CC disease.  
 XX  
 XX Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 375; DB 21; Length 1261;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-83;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CCTCTCCCTCCGAGCTGAGCTTACCTGGGCGCAACGAGCGAGCGGCGCGAGTGG 60  
 DB 1 CCTCTCCCTCCGAGCTGAGCTTACCTGGGCGCAACGAGCGAGCGGCGCGAGTGG 60  
 QY 61 AAGCTGAGTTCGGGGTGGGGCGGAGGCGGAGCTGCTGGTGGTGGAGCGGCGGAG 120  
 DB 61 AAGCTGAGTTCGGGGTGGGGCGGAGGCGGAGCTGCTGGTGGTGGAGCGGCGGAG 120  
 QY 121 GCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTGGCTTAG 180  
 DB 121 GCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTGGCTTAG 180  
 QY 181 GGGCCAAAGCCGCCACCGGCTCAAAAGCTCCAGGCGCTCCCGAGGACCGGGTGGCTCG 240  
 DB 181 GGGCCAAAGCCGCCACCGGCTCAAAAGCTCCAGGCGCTCCCGAGGACCGGGTGGCTCG 240  
 QY 241 GCGCTTCTTCGGTTCAGAAAGTCCGCTGGGGGCGAGTTCGTCGCAAGGGTTTCCTCG 300  
 DB 241 GCGCTTCTTCGGTTCAGAAAGTCCGCTGGGGGCGAGTTCGTCGCAAGGGTTTCCTCG 300  
 QY 301 AAAGAATCTGAGAGGGCGGCTGCTTGGACCGAGGGGAATCTCTGTGTAGCCTTGAAGC 360  
 DB 301 AAAGAATCTGAGAGGGCGGCTGCTTGGACCGAGGGGAATCTCTGTGTAGCCTTGAAGC 360  
 QY 361 GCGCGCGCCCGAGAG 375  
 DB 361 GCGCGCGCCCGAGAG 375  
 RESULT 2  
 ID ABA93727 standard; cDNA; 4052 BP.  
 XX  
 XX ABA93727;  
 AC ABA93727;  
 DT 30-APR-2002 (first entry)  
 XX  
 XX Human signal transduction cDNA clone amy2.2f18.  
 DE  
 XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;  
 KW gene therapy; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX

FN WO200198454-A2.  
 XX PD 27-DEC-2001.  
 XX  
 XX 25-APR-2001; 2001WO-IB02050.  
 PF  
 XX 25-APR-2000; 2000US-199380P.  
 PR  
 XX (GEHU-) GERMAN HUMAN GENOME PROJECT.  
 PA  
 XX Wiemann S;  
 PI  
 XX WPI; 2002-055860/07.  
 DR P-PSDB; ABB05689.  
 XX  
 XX Human cDNA sequences and clones derived from human fetal brain, fetal  
 PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic  
 PT screening and therapy -  
 XX  
 XX Claim 1; Page 174-175; 611pp; English.  
 XX  
 XX The present invention describes assemblages and computer readable media  
 CC comprising novel human cDNA sequences and clones derived from human  
 CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA  
 CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the  
 CC present invention which encode the proteins given in ABB05662 to  
 CC ABB05729. The human cDNA sequences and clones can be used in gene  
 CC therapy. The clones may be used in a variety of applications, for  
 CC example they may be used in profiling assays, for providing large arrays  
 CC of human genetic material for implementing large-scale screening  
 CC strategies and for treating diseases via gene therapy procedures.  
 XX  
 XX Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 other;  
 SQ  
 Query Match 98.6%; Score 369.8; DB 24; Length 4052;  
 Best Local Similarity 99.5%; Pred. No. 1.2e-81;  
 Matches 371; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 CTCCTTCCGAGCTGAGCTTACCTGGGCGCAACGAGCGAGCGGCGGCGAGTGGAA 62  
 DB 431 CTCCTTCCGAGCTGAGCTTACCTGGGCGCAACGAGCGAGCGGCGGCGAGTGGAA 490  
 QY 63 GCTGAGTTCGGGGTGGGGAGGCGAGCTGCTCCGTGGTGGTGGAGCGGCGAGAGC 122  
 DB 491 GCTGAGTTCGGGGTGGGGAGGCGAGCTGCTCCGTGGTGGTGGAGCGGCGAGAGC 550  
 QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTGGCTTAGG 182  
 DB 551 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTGGCTTAGG 610  
 QY 183 CCAGAAAGCCCGCCCGGCTCCAAAGCTCCAGGGGCTCCCGAGGCGACCGGTGCTCGGC 242  
 DB 611 CCAGAAAGCCCGCCCGGCTCCAAAGCTCCAGGGGCTCCCGAGGCGACCGGTGCTCGGC 670  
 QY 243 CTTCTCTCGGTGAGAAAGTCCGCCCTCGGGGCGAGTTCGTCGCAAGGGTTCTTCGAA 302  
 DB 671 CTTCTCTCGGTGAGAAAGTCCGCCCTCGGGGCGAGTTCGTCGCAAGGGTTCTTCGAA 730  
 QY 303 AGAATCTGAGAGGGCGGCTGCTTGGACCGAGGGGAATCTCTCTGTGTAGCCTTGGAAAGCCG 362  
 DB 731 AGAATCTGAGAGGGCGGCTGCTTGGACCGAGGGGAATCTCTCTGTGTAGCCTTGGAAAGCCG 790  
 QY 363 CCAGCCCGCAGAG 375  
 DB 791 CCAGCCCGCAGAG 803  
 RESULT 3  
 ID ABA98320 standard; cDNA; 978 BP.  
 XX  
 XX ABA98320;  
 AC  
 XX

DT 12-OCT-2001 (first entry)  
XX Human EST-derived coding sequence SEQ ID NO: 177.  
DE  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition; ss.  
XX Homo sapiens.  
OS  
XX WO200154477-A2.  
PN  
XX 02-AUG-2001.  
PD  
XX 25-JAN-2001; 2001WO-US02687.  
PF  
XX 25-JAN-2000; 2000US-0491404.  
PR  
XX 17-JUL-2000; 2000US-0617746.  
PR  
XX 03-AUG-2000; 2000US-0631451.  
PR  
XX 15-SEP-2000; 2000US-0663870.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
PI  
XX WPI; 2001-476164/51.  
DR  
XX P-PSDB; AM23661.  
DR  
XX  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
PT  
XX Claim 1; Page 299-300; 1275pp; English.  
PS  
XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
CC of the invention.  
XX  
XX Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 other;  
SQ  
Query Match 76.9%; Score 288.4; DB 22; Length 978;  
Best Local Similarity 99.7%; Pred. No. 1.1e-61;  
Matches 289; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 86 GAGGCGACTGTCCTGCTGAGCGCGCGGAGCGCGCGGCGGCGGCGGCTGATCGGCT 145  
Db 5 GTGGCGACTGTCCTGCTGAGCGCGCGGAGCGCGCGGCGGCGGCGGCTGATCGGCT 64  
QY 146 CCTCGAATCGGGAGGTCCTAGTGGGTCGCTTATAGGCGCCAAAGCCCCCAGCGGCTCCA 205  
Db 65 CCTCGAATCGGGAGGTCCTAGTGGGTCGCTTATAGGCGCCAAAGCCCCCAGCGGCTCCA 124  
QY 206 AAGCTCCAGGCGCTCCCGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 265  
Db 125 AAGCTCCAGGCGCTCCCGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 184  
QY 266 CCCCTGGGGGCAAGTTCCTCCCAAGGGTTTCTCGAAGAGATCTGAGAGGCGCGAGTCT 325  
Db 185 CCCCTGGGGGCAAGTTCCTCCCAAGGGTTTCTCGAAGAGATCTGAGAGGCGCGAGTCT 244  
QY 326 TGACCGAGGGAATCTCTGTGTAGACCTTGAAGCGCGGCGGCGGCGGCGGCGGCTCA 375  
Db 245 TGACCGAGGGAATCTCTGTGTAGACCTTGAAGCGCGGCGGCGGCGGCGGCGGCTCA 294  
RESULT 4  
AAK52345

AAK52345 standard; cDNA; 1045 BP.  
AAK52345;  
06-NOV-2001 (first entry)  
Human polynucleotide SEQ ID NO 890.  
Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.  
XX Homo sapiens.  
OS  
XX WO200157190-A2.  
PN  
XX 09-AUG-2001.  
PD  
XX 05-FEB-2001; 2001WO-US04098.  
PF  
XX 03-FEB-2000; 2000US-0496914.  
PR  
XX 27-APR-2000; 2000US-0560875.  
PR  
XX 20-JUN-2000; 2000US-0598075.  
PR  
XX 19-JUL-2000; 2000US-0620325.  
PR  
XX 01-SEP-2000; 2000US-0654936.  
PR  
XX 15-SEP-2000; 2000US-0663561.  
PR  
XX 20-OCT-2000; 2000US-0693325.  
PR  
XX 30-NOV-2000; 2000US-0728422.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
DR  
XX P-PSDB; AM79212.  
DR  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
PT  
XX Claim 1; Page 2934-2935; 6221pp; English.  
PS  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAK80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX  
SQ Sequence 1045 BP; 222 A; 289 C; 314 G; 220 T; 0 other;  
Query Match 76.5%; Score 286.8; DB 22; Length 1045;  
Best Local Similarity 99.3%; Pred. No. 2.7e-61;  
Matches 288; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 86 GAGGCGACTGTCCTGCTGAGCGCGCGGAGCGCGGCGGCGGCGGCTGATCGGCT 145  
Db 12 GTGGCGACTGTCCTGCTGAGCGCGCGGAGCGGCGGCGGCGGCGGCTGATCGGCT 71  
QY 146 CCTCGAATCGGGAGGTCCTAGTGGGTCGCTTATAGGCGCCAAAGCCCCCAGCGGCTCCA 205  
Db 72 CCTCGAATCGGGAGGTCCTAGTGGGTCGCTTATAGGCGCCAAAGCCCCCAGCGGCTCCA 131

QY 206 AAAGTCCAGGCTCCCGAGGACCGGTGCTTGGCCCTTCTTGGTCTGAGAAAGTGC 265  
DB 132 AAAGTCCAGGCTCCCGAGGACCGGTGCTTGGCCCTTCTTGGTCTGAGAAAGTGC 191  
QY 266 CCCCTGGGGGAGTTCGTCCTCCAAAGGTTTCTCCAAAGATCTGAGAGGCGCAGTCT 325  
DB 192 CCCCTGGGGGAGTTCGTCCTCCAAAGGTTTCTCCAAAGATCTGAGAGGCGCAGTCT 251  
QY 326 TGACCGAGGGAATCTCTGTAGCTTGGAGCGCGAGCCCGCAGAAAG 375  
DB 252 TGACCGAGGGAATCTCTGTAGCTTGGAGCGCGAGCCCGCAGAAAG 301

RESULT 5  
ID AAZ40538  
XX AAZ40538 standard; DNA; 2806 BP.  
AC AAZ40538;  
XX  
DT 18-FEB-2000 (first entry)  
XX  
DE Full length human PAK5 DNA.  
KW Antirheumatic; antiarthritic; antiinflammatory; antiallergic; osteopathic;  
KW antipsoriatic; antiarteriosclerotic; antiasthmatic; immunosuppressive;  
KW neuroprotective; cardiac; cerebroprotective; cytostatic; antidiabetic;  
KW vulnery; STE20; protein kinase; STIK2; STIK3; STIK4; STIK5; STIK6; STIK7;  
KW ZC1; ZC2; ZC3; ZC4; KHS2; SULU1; SULU3; GEX2; PAK4; PAK5; antagonist;  
KW antibody; gene therapy; rheumatoid arthritis; atherosclerosis; asthma;  
KW inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;  
KW rhinitis; autoimmunity; organ transplantation; multiple sclerosis;  
KW myocardial infarction; cardiovascular disease; stroke; renal failure;  
KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;  
KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;  
KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;  
KW mesangial disorder; growth regulation; wound healing; T cell activation;  
KW immunosuppressant; ss.  
XX  
OS Homo sapiens.

XX WO9953036-A2.  
PN  
PD 21-OCT-1999.  
XX  
PF 13-APR-1999; 99WO-US08150.  
XX  
PR 14-APR-1998; 98US-0081784.  
XX  
PA (SUGEN-) SUGEN INC.  
XX  
PI Plowman G, Martinez R, Whyte D;  
XX  
XX WPI; 1999-611301/S2.  
DR P-PSDB; AAY55964.  
XX  
PT Novel kinase-related polypeptides used for the diagnosis and treatment  
PT of kinase-related diseases and disorders.  
XX  
PS Disclosure; Page 365; 387pp; English.

XX This sequence represents the coding sequence for a novel STE20-related  
XX protein kinase. The invention relates to nucleic acid molecule encoding  
XX a kinase polypeptide selected from STIK2, STIK3, STIK4, STIK5, STIK6,  
XX STIK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEX2, PAK4 and PAK5. The  
XX proteins are used to identify agonists and antagonists, and to raise  
XX antibodies. The polynucleotides are useful in gene therapy protocols. The  
XX polynucleotides, polypeptides, antibodies, antagonists and agonists may  
XX be used to treat diseases such as immune-related disorders and diseases  
XX (e.g. rheumatoid arthritis, atherosclerosis, chronic inflammatory bowel  
XX disease (e.g. Crohn's disease), asthma, osteoarthritis, psoriasis,  
XX atherosclerosis, rhinitis, autoimmunity, and organ transplantation,  
XX chronic inflammatory pelvic disease, multiple sclerosis, organ  
XX transplantation, myocardial infarction, cardiovascular disease, stroke,

CC renal failure, oxidative stress-related neurodegenerative disorders (e.g.  
CC amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome),  
CC cancer, cardiomyopathies, ischemic disorders, inflammatory disorders,  
CC diabetes mellitus, fibrotic and mesangial disorders. The proteins may  
CC also be useful for cell growth regulation (e.g. in wound healing), T cell  
CC activation, mitosis control, and as immunosuppressants.  
XX  
SQ Sequence 2806 BP; 523 A; 957 C; 869 G; 457 T; 0 other;

Query Match 10.7%; Score 40.2; DB 20; Length 2806;  
Best Local Similarity 49.8%; Pred. No. 1.3;  
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 23 ACCCTGGGCGCAACGAGCGAGCGGCGCGAGTGGAGTCTGAGTCCGGGGTGGC 82  
DB 567 ACCACGCCAGAGGGGGGCCAGGGAAGCGAGCGCCGGTTCGGCGGTACAGC 626  
QY 83 GGGGAGGCGACTGTCCGTGTGCTGTAGCGCGCGGCGAGCGGGCGCGAGCGGTGATCG 142  
DB 627 GAGCGGGTGGCGGCGAGTGTGACAGCGGCGGCGGCGGCGAGAGGCCCAAGTCT 686  
QY 143 GCTCCCTCGAACTGGGAGGTCCAGTGGGTGCGTCTAGGCGCCCAAGCCCCCGGCT 202  
DB 687 TCCAGGAGGGGCTCAGGGGGTCCCGAGGAGTCTCCCGGGACAAACGCCCTCTCCGGG 746  
QY 203 CCAAAAGCTCCAGCGGCTCCCGCAG 227  
DB 747 CCTGATGTGCGCACCCCGCCAGCTG 771

RESULT 6  
AAS83495  
ID AAS83495 standard; cDNA; 2811 BP.  
XX  
AC AAS83495;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #19299.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.

XX WO200175067-A2.  
PN  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR P-PSDB; ABG19308.

XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
PS Claim 1; SEQ ID No 19299; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags







CC differentially expressed gene product in a test sample from a cell  
CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.  
XX  
SQ Sequence 1337 BP; 42 A; 577 C; 27 G; 22 T; 669 other;

Query Match 10.3%; Score 38.6; DB 20; Length 1337;  
Best Local Similarity 41.5%; Pred. No. 3;  
Matches 71; Conservative 0; Mismatches 100; Indels 0; Gaps 0;  
QY 12 GAGCTGAGCTTACCTCGGCGCAACAGCAGCGAGCGGCGGCGAGTGAAGCTGAGTT 71  
Db 697 GGGGNGGNGNGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGG 638  
QY 72 CCGGGTGGGGGAGGAGCGACTGTCCTGCTGAGCGCGCGGCGAGAGCGGCGCGGA 131  
Db 637 GCGGGGCGNGGNGGCGGNGGNGGNGGNGGCGGCGGNGGNGGNGGNGGNGGNG 578  
QY 132 GGGGCTGATCGCTCCCTCGAACTGGGGAGGTCCTGAGTGGGTCGCTTAGGG 182  
Db 577 GGGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGG 527

RESULT 12  
AA06422  
ID AAD06422 standard; DNA; 1776 BP.  
XX AC AAD06422;  
XX  
XX 10-AUG-2001 (first entry)  
XX Human p21-activated kinase 4 (PAK4) DNA coding sequence.  
XX  
XX Human, p21-activated kinase 4; PAK4; genetic mapping; neuroprotective;  
XX immunosuppressive; gene therapy; cancer; angiogenesis-related disorder;  
XX serine-threonine kinase; restriction fragment length polymorphism; RFLP;  
XX central nervous system disorder; immune-related disorder; cytostatic; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200136602-A2.  
XX  
XX 25-MAY-2001.  
XX  
XX 14-NOV-2000; 2000WO-BE10736.  
XX  
XX 15-NOV-1999; 99US-0439756.  
XX  
XX (PHAA ) PHARMACIA & UPJOHN SPA.  
XX  
XX Dahlberg M, Moll J, Galvani A;  
XX  
XX WPI; 2001-355626/37.  
XX  
XX New nucleic acid encoding p21-activated kinase 5 protein for inducing  
XX an immune response, treating cancers, angiogenesis-related disorders,  
XX disorders of the central nervous system and immune-related disorders -  
XX  
XX Example 1; Page 36-37; 76pp; English.

XX The present DNA sequence is the coding sequence (CDS) of human p21-  
CC activated kinase 4 (PAK4) DNA. The protein encoded by PAK1 DNA is 85%  
CC identical to PAK5 protein which is a serine-threonine kinase. PAK5 is  
CC useful for inducing an immune response in a mammal against PAK5  
CC polypeptide. PAK5 proteins are useful in regulating cell proliferation,  
CC cell migration, cell differentiation, cytoskeletal organisation, gene  
CC expression, cell cycle progression, and cell death. PAK5, is useful in  
CC the search for novel agents that can modify and/or control the above  
CC processes. PAK5 DNA is useful for screening restriction fragment length  
CC polymorphism (RFLP) associated with certain disorders, as well as for  
CC genetic mapping. PAK5 DNA is also useful as diagnostic tools for probing  
CC gene expression in various tissues. PAK5 polypeptides are also useful as  
CC antigens for raising antibodies, and for screening compounds that  
CC modulate the activity of PAK5. PAK5 is also useful in pharmaceutical  
CC compositions, and in the manufacture of medicaments for treating diseases  
CC such as cancers, angiogenesis-related disease, diseases of the central  
CC nervous system and diseases due to inappropriate activation of immune  
CC responses. PAK5 DNA is also useful in gene therapy.  
XX  
SQ Sequence 1776 BP; 337 A; 616 C; 568 G; 255 T; 0 other;  
Query Match 10.3%; Score 38.6; DB 22; Length 1776;  
Best Local Similarity 49.3%; Pred. No. 3.1;  
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 23 ACCCTGGCGCAACAGCAGCGAGCGGCGGCGGCGAGTGAAGTTCGGGGTGGGC 82  
Db 367 ACCACGGCGAGAGGGGCGCCAGGAGCAGCAGCGAGCGGCGGTCACAGC 426  
QY 83 GGGGAGGCGACTGTCCGTGGTCTGAGCGCGCGGAGAGCGGCGGCGGCTGATCG 142  
Db 427 GAGGAGGTGGCGCGAGTGTGTACAGGCGACCGCGGCGGCGAGAGGCGCCCAAGTCT 486  
QY 143 GCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCTGCTTAGGGCCCAAGCCCCCGGCT 202  
Db 487 TCCAGGAGGGCTCAGGGGGTCCCGAGGAGTCTCCCGGAGCAACAGCCCTCTCCGGG 546  
QY 203 CCAGAGCTCCAGGGGCTCCCGAG 227  
Db 547 CCTGATGTCGCGACCCCCCAGCCTG 571

RESULT 13  
AAZ40657  
ID AAZ40657 standard; DNA; 1941 BP.  
XX AC AAZ40657;  
XX  
XX 08-MAR-2000 (first entry)  
XX  
XX Human serine/threonine kinase, PAK4 encoding DNA.  
XX  
XX PAK4; serine/threonine kinase; GTPase; intracellular signal cascade;  
XX Rac; Cdc42H; morphogenesis; mitogenesis; JNK; p38 MAP kinase; human;  
XX actin polymerization; filopodia; cancer; arthritis; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO9963073-A1.  
XX  
XX 09-DEC-1999.  
XX  
XX 21-MAY-1999; 99WO-US11341.  
XX  
XX 21-MAY-1998; 98US-0082737.  
XX  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX  
XX Minden A;  
XX  
XX WPI; 2000-072881/06.  
XX  
XX P-PSDB; AAY59128.





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:08:07 ; Search time 1108.64 Seconds  
(without alignments)  
8221.015 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_1\_375

Perfect score: 375

Sequence: 1 cctcccttcgagctgagc.....gaagcgccagccccagaag 375

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*  
2: em\_esthum:\*\*  
3: em\_estin:\*\*  
4: em\_estmu:\*\*  
5: em\_estov:\*\*  
6: em\_estpl:\*\*  
7: em\_estro:\*\*  
8: em\_estc:\*\*  
9: gb\_est1:\*\*  
10: gb\_est2:\*\*  
11: gb\_est3:\*\*  
12: gb\_est4:\*\*  
13: gb\_est5:\*\*  
14: em\_estfun:\*\*  
15: em\_estom:\*\*  
16: em\_gss\_hum:\*\*  
17: em\_gss\_inv:\*\*  
18: em\_gss\_pln:\*\*  
19: em\_gss\_vrt:\*\*  
20: em\_gss\_fun:\*\*  
21: em\_gss\_mam:\*\*  
22: em\_gss\_mus:\*\*  
23: em\_gss\_pro:\*\*  
24: em\_gss\_rnd:\*\*  
25: em\_gss\_rnd:\*\*  
26: em\_gss\_png:\*\*  
27: em\_gss\_vrt:\*\*  
28: gb\_gss1:\*\*  
29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	346	92.3	1201	9	AL534136
2	213.8	57.0	972	13	BQ066875
3	59.4	15.8	1103	13	BX403654
4	59	15.7	935	29	CNS006XK

C	5	53.8	14.3	1101	29	CNS00397	AL063912 Drosophil
C	6	53.6	14.3	982	13	BX415111	BX415111
C	7	53.4	14.2	1000	13	BX407619	BX407619
C	8	53	14.1	693	29	AG060136	AG060136 Pan trogl
C	9	52.8	14.1	1201	9	AL538546	AL538546
C	10	51.8	13.8	935	29	CNS006XK	AL066051 Drosophil
C	11	51.6	13.8	925	29	CNS0091P	AL033013 Drosophil
C	12	51.2	13.7	1103	13	BX403654	BX403654
C	13	50.6	13.5	804	13	BX425573	BX425573
C	14	50.6	13.5	925	29	CNS0091P	AL053013 Drosophil
C	15	50.6	13.5	1198	13	BX371097	BX371097
C	16	50	13.3	964	29	AG127812	AG127812 Pan trogl
C	17	49.8	13.3	888	12	BG809572	BG809572 mgct001xa
C	18	49.6	13.2	932	29	CNS0072Q	AL066742 Drosophil
C	19	49.4	13.2	1269	12	BM807701	BM807701 AGENCOURT
C	20	49.2	13.1	1061	13	BX391246	BX391246
C	21	48.8	13.0	863	29	AG132157	AG132157 Pan trogl
C	22	48.8	13.0	1144	13	BX415926	BX415926
C	23	48.6	13.0	970	13	BUS38852	BUS38852 AGENCOURT
C	24	48.6	13.0	1000	13	BX407619	BX407619
C	25	48.6	13.0	1009	29	CNS010EW	AL098882 Drosophil
C	26	48.4	12.9	515	13	BX424977	BX424977
C	27	48.4	12.9	1179	13	BX425030	BX425030
C	28	48.2	12.9	956	13	BUS01227	BUS01227 AGENCOURT
C	29	48.2	12.9	966	13	BQ672441	BQ672441 AGENCOURT
C	30	48.2	12.9	1203	29	CNS015Y4	AL106054 Drosophil
C	31	48	12.8	1171	12	BM914930	BM914930 AGENCOURT
C	32	47.8	12.7	873	29	AG043471	AG043471 Pan trogl
C	33	47.8	12.7	897	29	CNS015LJ	AL105601 Drosophil
C	34	47.6	12.7	559	13	BX384405	BX384405
C	35	47.4	12.6	1057	13	BX349688	BX349688
C	36	47.2	12.6	377	10	BG361509	BG361509 gb60d11.y
C	37	47.2	12.6	732	28	AZ196033	AZ196033 SP.1031.A
C	38	47.2	12.6	1201	9	AL581390	AL581390
C	39	46.8	12.5	936	12	BG852371	BG852371 1024034A0
C	40	46.8	12.5	1048	13	BQ073800	BQ073800 AGENCOURT
C	41	46.8	12.5	1198	13	BX371097	BX371097
C	42	46.6	12.4	442	28	BH172722	BH172722 SALK_0060
C	43	46.6	12.4	860	29	CNS0226P	AL220570 Tetraodon
C	44	46.6	12.4	1136	13	BX446909	BX446909
C	45	46.4	12.4	626	29	CNS022N9	AL221166 Tetraodon

#### ALIGNMENTS

RESULT 1  
AL534136  
LOCUS  
DEFINITION  
AL534136 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
CS0DF003Y102 5-PRIME, mRNA sequence.  
ACCESSION  
AL534136  
VERSION  
AL534136.2 GI:30539643  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1201)  
Li W.B., Gruber C., Jessee J. and Polayes D.  
Full-length cDNA libraries and normalization  
Unpublished  
On Feb 13, 2001 this sequence version replaced gi:12797629.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6147.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DF003Y102&cluster=6147.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DF005BE01QP1.

FEATURES source

```

1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0F005102"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. library was not normalized."
331 c 402 g 214 t 33 others

```

221 a 331 c 402 g 214 t 33 others

ORIGIN

Query Match	92.3%; Score 346; DB 9; Length 1201;
Best Local Similarity	98.9%; Pred. No. 1.3e-69;
Matches 369; Conservative 0; Mismatches 1; Indels 3; Gaps 2;	
3	CTCCCTTCCGAGCTGAGCTTTACCTCGGCGCAAAACGAGCGAGGCGAGGCGCGAGTGCAG 62
459	CTCCTTTCGAGCTGAGCTTTACCTCGGCGCAAAACGAGCGAGGCGAGGCGCGAGTGCAG 518
63	GCTCGAGTTCCGGGTTGGGCGGGAGGGCGACTGTCCGTGTGTCGTAGCGCGCGGAGAGC 122
519	GCTCGAGTTCCGGGTTGGGCGGGAGGGCGACTGTCCGTGTGTCGTAGCGCGCGGAGAGC 578
123	GGGCGGAGGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGGTGCCTTAGGG 182
579	GGGCGCGNA--GGCTGATGGCTCCCTCGAACTGGGAGGTCAGTGGGGTGCCTTAGGG 636
183	CCCAAGCCCCCACCCGGCTCCAAAAGCTCCAGGCGCTCCCGAGGCAACGGTGCCTCGC 242
637	CCCAAGCCCCCACCCGGCTCCAAAAGCTCCAGGCGCTCCCGAGGCAACGGTGCCTCGC 696
243	CTTTCCTTCGGTCAGAAAGTCGCCCTCGGGGCGAGTTCGTCCCAAAGGGTTTCCTCGAA 302
697	CTTTCCTTCGGTCAGAAAGTCGCCCTCGGGGCGAGTTCGTCCCAAAGGGTTTCCTCGAA 756
303	AGATCTGAGAGGCGCAGTCCTTTGACCGAGGGGAATCTCTCTGTGTAGCCTTGAAGCG 362
757	AGATCTGAGAGGCGGCA--TCTTTGACCGAGGGGAATCTCTCTGTGTAGCCTTGAAGCG 815
363	CCAGCCCCAGAAG 375
816	CCAGCCCCAGAAG 828

RESULT 2  
 BQ066875  
 LOCUS  
 DEFINITION  
 5', mRNA sequence.  
 ACCESSION  
 BQ066875  
 VERSION  
 BQ066875.1  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 972)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 AUTHORS  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE  
 Unpublished  
 JOURNAL  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapsf@mail.nih.gov](mailto:cgapsf@mail.nih.gov)  
 COMMENT  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

## FEATURES

```

1. .972
/organism="Homo sapiens"
/mol type="cDNA"
/db xref="taxon:9606"
/clone="IMAGE:5762608"
/lab host="DH10B"
/clone lib="N1H MGC 114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Intivrogen). Research Genetics tracking code 019. Note:
this is a N1H MGC Library."
```

BASE COUNT	ORIGIN
159 a	343 c 320 g 150 t

## ORIGIN

Query Match	57.0%;	Score 213.8;	DB 13;	Length 972;
Seet Local Similarity	90.7%;	Pred. No. 3.1e-39;		
Matches 262;	Conservative 0;	Mismatch 22;	Indels 5;	Gaps 3;
3	CTCCCTTCGGAGCTGAGCTTACCTCGGGCGCAAAAGAGCGAGCGAGGGCGCGAGTGGAA	62		
412	CTCCCTTCGGAGCTGAGCTTACCTCGGGCGCAAAAGAGCGAGCGAGGGCGCGAGTGGAA	471		
63	GCTGAGTTCCGGGTGGCGGGGAGCGACTGTCCGTGGTCTGTAGCGCCCGCGAGAGC	122		
472	GCTGAGTTCCGGGTGGCGGGGAGCGACTGTCCGTGGTCTGTAGCGCCCGCGAGAGC	531		
123	GGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTGCCTTAGGG	182		
532	GGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGGCCGTTGGGTGCCTTAAGG	591		
183	CCCAAGCGCCCAACCCCGGCTCCAAAGCTCCCA-GGGCCTCCCGAGGACACGGTG--CTC	239		
532	GCCAAGCGCCCAACCCCGGCTCCAAAGCTCCCA-CGGCCCTCCCCAGGAGACGGGGGCTTC	651		
240	GGCCCTTCCTT--CGGTCAAGAGTCGCCCTCGGGGCGAGTTCGTCC	286		
652	GGCCCTTCCTTTCGGGACAAAGTCCTCCGCCCTCGGGGGCGCCGCC	700		

RESULT_3	
EX403654	
LOCUS	mRNA linear EST 15-MAY-2003
DEFINITION	EX403654 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA007ZH12 3-PRIME, mRNA sequence.
ACCESSION	EX403654
VERSION	EX403654.1 GI:30762430
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1103) <i>Li,W.B., Gruber,C., Jessee,J. and Polayes,D.</i> Full-length cDNA libraries and normalization Unpublished
AUTHORS	Contact: Genoscope
TITLE	Genoscope - Centre National de Sequencage
JOURNAL	Bp 191 91006 EVRY cedex - France
COMMENT	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of







Qy	28	GGCGGCAACGACGCAGGCGAGGGCCGCGAGTCTGGAGTTCCGGGTGGCGGGGA	87
Db	669	GSGGGCGSGCAGCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCGSCG	728
Qy	88	GCGACTGTCCGTGGTGCTGAGCGCGCGGCGAGCGCGCGAGCGGCTGATCGGCTCC	147
Db	729	CSSGSGCCSCCGCSUSSSBSCCGCGCMSCGCGSSCGCGCGCGCGCGSCG	788
Qy	148	CTCGAACTGGGAGGTCCTAGTGGGTGCTTAGGGCCCCAAGCCCCACC CGGCTCCAAA	207
Db	789	CSGGCCSGCGCGCGSGSCGCGCGSSSSGSGSCGSCGCGCGSCCSGGGCGCSSSC	848
Qy	208	AGTCCAGGGCCTCCCAGGACA CCGGTGCTCGGCCCTTCTTCGTCA GAAAGTCGCC	267
Db	849	SGCGCGSGGSSCCGSGCGCGSCGSGSGSGSSGSSGSGCGSGCGSGCGSGGSGCS	908
Qy	268	CCTGGGGCAGTTCGTCCC	286
Db	909	CSCGCGCGCSGSSCSCC	927

LOCUS	CNS0091P	925 bp	DNA	linear	GSS 03-JUN-1999			
DEFINITION	Drosophila melanogaster; genome survey sequence HET3 end of BAC # BACR19D16 of RGC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.							
ACCESSION	AL053013							
VERSION	AL053013.1	GI:4934461						
KEYWORDS	GSS.							
SOURCE	Drosophila melanogaster (fruit fly)							
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 925)							
REFERENCE	Genoscope.							
AUTHORS	Direct Submission							
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :							
JOURNAL	BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr							
COMMENT	- Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).							

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mamooser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial *NotI* digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source  
1. . 925  
Location/Qualifiers

```

source
i. 1943
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon.7227"
/clone="BACRI9D16"
/clone_lib="RPCT-98"
/note="end : TET3"

BASE COUNT      120 a      61 c      172 t      511 others
ORIGIN

Query Match      13.8%; Score 51.6; DB 29; Length 925;
Best Local Similarity 13.4%; Pred. No. 0.063;
Matches 37; Conservative 136; Mismatches 104; Indels 0; Gaps 0;

Qv 2 CCTCCCTTCGGAGCTGAGCTTACCTCTGGGCGCAACGAGCGAGCGGGCGCGAGTGGG 61

```







Db 488 TCGGAGGAGCGCGGCTCGACGCGGTGCGCGGAGTGATGACAGCGCC 547  
QY 193 CCACCGGCTCCAAAGCTCCACAGGCTCCCGAGGACCGGTCGCGCCCTTCCTCG 252  
Db 548 CCGGCGGACCGCGAGCGCGCGGCTCCCGACCGCGCGGCTCCCGCGG 607

## RESULT 2

US-09-082-737-1  
; Sequence 1, Application US/09082737  
; Patent No. 6013500  
; GENERAL INFORMATION:  
; APPLICANT: Minden, Audrey  
; TITLE OF INVENTION: FAK3; A No. 6013500el Gene Encoding A Serine/  
; TITLE OF INVENTION: Threonine Kinase  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 11230  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,737  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/55311  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1941 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 166..1938  
US-09-082-737-1

Query Match 10.3%; Score 38.6; DB 3; Length 1941;  
Best Local Similarity 49.3%; Pred. No. 0.43;  
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 23 ACCCTGGCGCAACGAGCGAGGCGCGGCGCGAGTGGAGCTTCCGGGCTGGCG 82  
Db 532 ACCAGCGCAGAGCGGCGCCAGGAGGCGAGGCGCGGTTCCGCGCTCACAGC 591  
QY 83 GGGGAGCGACTGTCCGTGTGTGTGAGCGCGCGGAGAGGCGCGGCGGCTGATCG 142  
Db 592 GAGCGAGTGGCGGCGAGTGTGACAGCGCGAGCGGCGGCGGCGGCGGCAAGTCT 651  
QY 143 GCTCCCTCGAACTGGGGAGGTCCAGTGGGCTGCGTTAGGGCCCAAGCCCGCGGCT 202  
Db 652 TCCAGGAGGAGGCTCAGGGGGTCCCGAGGAGTCTCCCGGAGCAAGCCCTCTCGGG 711  
QY 203 CCAAAAGCTCCCGAGGCTCCCGCAG 227  
Db 712 CCTGATGTCGCGACCCCGCCAGCCTG 736

## RESULT 3

US-09-128-155-16/c  
; Sequence 16, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 152331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(152331)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-16

Query Match 9.9%; Score 37.2; DB 3; Length 152331;  
Best Local Similarity 53.4%; Pred. No. 1.6;  
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 38 GAGCGAGCGAGCGCGGCGGAGTGGAGTTCGCGGCGGCGGAGCGGAGCTGTC 97  
Db 22281 GCGGCGGCGCGGCGGCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 22222  
QY 98 CCGTGTGTGAGCGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 157  
Db 22221 GGTGGGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 22162  
QY 158 GAGGTCCTCAGTGGGCTGCTTAGGGC 183  
Db 22161 GCGGCGGCTCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 22136

## RESULT 4

US-09-904-615-16/c  
; Sequence 16, Application US/09904615  
; Patent No. 6566325  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032P1  
; CURRENT APPLICATION NUMBER: US/09/904,615  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 09/511,554  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/097,917  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 60/098,634  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 2608  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-904-615-16

Query Match 9.7%; Score 36.2; DB 4; Length 2608;  
Best Local Similarity 54.0%; Pred. No. 1.8;  
Matches 74; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 26 CTGGCGCAACGAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 85

Db 1824 CAGGCTCACAGCACCAGGATCGGAGCGTGGCCAGTCTGGGCCCTGGCTGTGC 1765  
QY 86 GAGCGAGTCTCCCTGGTCTGAGCGCGGCGAGCGGCGCGGAGCGGCTGATCGCT 145  
Db 1764 GCGCGGCTGTCTGGGCTCTGCTCAGCGGCGAGAGCGGTCCAGAGCAGT 1705  
QY 146 CCCTCGAACTGGGAGG 162  
Db 1704 CCCAGCAGGAGCGG 1688

## RESULT 5

US-09-320-878-19/c

; Sequence 19, Application US/09320878A

; Patent No. 6117659

; GENERAL INFORMATION:

; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Melanie C.

; APPLICANT: BETLACH, Mary C.

; APPLICANT: MCDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

; FILE REFERENCE: 300622002120

; CURRENT APPLICATION NUMBER: US/09/320,878A

; CURRENT FILING DATE: 1999-05-27

; EARLIER APPLICATION NUMBER: CIP OF 09/141,908

; EARLIER FILING DATE: 1998-08-28

; EARLIER APPLICATION NUMBER: CIP OF 09/073,538

; EARLIER FILING DATE: 1998-05-06

; EARLIER APPLICATION NUMBER: CIP OF 08/846,247

; EARLIER FILING DATE: 1997-04-30

; EARLIER APPLICATION NUMBER: 60/119,139

; EARLIER FILING DATE: 1999-02-08

; EARLIER APPLICATION NUMBER: 60/100,880

; EARLIER FILING DATE: 1998-09-22

; EARLIER APPLICATION NUMBER: 60/087,080

; EARLIER FILING DATE: 1998-05-28

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 38506

; TYPE: DNA

; ORGANISM: Streptomyces venezuelae

US-09-320-878-19

Query Match 9.7%; Score 36.2; DB 3; Length 38506;  
Best Local Similarity 45.6%; Pred. No. 2.5;  
Matches 128; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 93 CTGTCCGTGGTCTGAGCGCGGCGAGAGCGGCGGCGGCTGATCGGCTCCCTCGA 152  
Db 3393 CTGGCGGAGTACGACGTGCGGGGTCTCGGATCCGGGTGTACAGCCCTCCACGTC 3334  
QY 153 ACTGGGAGTCCAGTGGGTCTGTAGGGCCCCAAAGCCCCCGGCTCCAAAGCTC 212  
Db 3333 CCAGCCGGTCTCCGGGAACTCCGAGATCGGCTCCCGCGCGCCACAGCCGCCA 3274  
QY 213 CCAGGGCTCCCGAGCGACCGGTCTCGGCCCTTCTTCGGTCAAGAGTCCGCCCTGG 272  
Db 3273 CAGGTCTCCGGGAGCGACCCCGGCGAGCGGCGGCGGCGGATGCCAGATCGCCAC 3214  
QY 273 GGGCAGTTCGTCCTCCAAAGGTTCTTCGAAAGATCTGAGAGGCGCGAGTCTTTGACCGA 332  
Db 3213 CGGCTCGTCAAGCGACCGCGCGCGCCCGGCGGCTCCGACCGCGGCTTCGC 3154  
QY 333 GGGATCTCTCTGTGTAGCTTGGAGCCGCGCGCCCGCA 373  
Db 3153 CGCCGCTCCCGTGCACGAGGAGCGTGTCTCCGCA 3113

## RESULT 6

US-09-141-908-1/c

; Sequence 1, Application US/09141908  
; Patent No. 6503741  
; GENERAL INFORMATION:  
; APPLICANT: ASHLEY, Gary  
; APPLICANT: BETLACH, Melanie C.  
; APPLICANT: BETLACH, Mary  
; APPLICANT: MCDANIEL, Robert  
; APPLICANT: TANG, Li  
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a  
; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold  
; FILE REFERENCE: 300622002100  
; CURRENT APPLICATION NUMBER: US/09/141,908  
; CURRENT FILING DATE: 1998-08-28  
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538  
; EARLIER FILING DATE: 1998-05-06  
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
; EARLIER FILING DATE: 1997-04-30  
; EARLIER APPLICATION NUMBER: PROV. 60/076,919  
; EARLIER FILING DATE: 1998-03-05  
; EARLIER APPLICATION NUMBER: PROV. 60/087,080  
; EARLIER FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 38506  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-141-908-1

Query Match 9.7%; Score 36.2; DB 4; Length 38506;  
Best Local Similarity 45.6%; Pred. No. 2.5;  
Matches 128; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 93 CTGTCCGTGGTCTGAGCGCGGCGAGAGCGGCGGCGGCTGATCGGCTCCCTCGA 152  
Db 3393 CTGGCGGAGTACGACGTGCGGGGTCTCGGATCCGGGTGTACAGCCCTCCACGTC 3334  
QY 153 ACTGGGAGTCCAGTGGGTCTGTAGGGCCCCAAAGCCCCCGGCTCCAAAGCTC 212  
Db 3333 CCAGCCGGTCTCCGGGAACTCCGAGATCGGCTCCCGCGCGCCACAGCCGCCA 3274  
QY 213 CCAGGGCTCCCGAGCGACCGGTCTCGGCCCTTCTTCGGTCAAGAGTCCGCCCTGG 272  
Db 3273 CAGGTCTCCGGGAGCGACCCCGGCGAGCGGCGGCGGATGCCAGATCGCCAC 3214  
QY 273 GGGCAGTTCGTCCTCCAAAGGTTCTTCGAAAGATCTGAGAGGCGCGAGTCTTTGACCGA 332  
Db 3213 CGGCTCGTCAAGCGACCGCGCGCGCCCGGCGGCTCCGACCGCGGCTTCGC 3154  
QY 333 GGGATCTCTCTGTGTAGCTTGGAGCCGCGCGCCCGCA 373  
Db 3153 CGCCGCTCCCGTGCACGAGGAGCGTGTCTCCGCA 3113

## RESULT 7

US-09-657-440-19/c

; Sequence 19, Application US/09657440

; Patent No. 6509455

; GENERAL INFORMATION:

; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Melanie C.

; APPLICANT: BETLACH, Mary C.

; APPLICANT: MCDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

; FILE REFERENCE: 300622002120

; CURRENT APPLICATION NUMBER: US/09/657,440

; CURRENT FILING DATE: 2000-09-07

; PRIOR APPLICATION NUMBER: 09/320,878

; PRIOR FILING DATE: 1999-05-27

; PRIOR APPLICATION NUMBER: CIP OF 09/141,908

; PRIOR FILING DATE: 1998-08-28

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 38506

; TYPE: DNA

; ORGANISM: Streptomyces venezuelae

US-09-657-440-19

Query Match 9.7%; Score 36.2; DB 4; Length 38506;  
Best Local Similarity 45.6%; Pred. No. 2.5;  
Matches 128; Conservative 0; Mismatches 153; Indels 0; Gaps 0;  
QY 93 CTGTCGCTGCTCTGAGCGCGGCGAGAGCGGCGGCTGATCGGCTCCCTCGA 152  
DB 3393 CTGGCGGACGTACGACGTCGCCGGGTGCTCGGATCGGCTGACAGCCCTCCACGTC 3334  
QY 153 ACTGGGAGTCTCAGTGGGTCGCTTAGGGCCCAAGCCCCACCGGCTCCAAAAGTCT 212  
DB 3333 CCAGCGCGGTCTCGCGGAACCTCCAGATCGCTCCCGCGCCGACCGCGCCA 3274  
QY 213 CCAGGGCTCCCGAGCACCGGTGCTCGGCCCTTCTTCGTCAGAAAGTCGCCCTGG 272  
DB 3273 CAGGTCTCTCGCGAGGACCCACCGGCGAGGCGGCGCATGCGACGATCGCCAC 3214  
QY 273 GGGCAGTTCGTCGCAAGGCTTCTCTCGAAGAACTGAGAGGCGCGCAGTCCTTGACCGA 332  
DB 3213 CGGCTGCTGACGGCACCGCGCGCGCCACCGAGCGGCTCCGACCGCGGTTGCG 3154  
QY 333 GGAATCTCTGTGTAGCTTTGGAAGCGCGCCAGCCCGCA 373  
DB 3153 CGCGGCTCCCGTGCACGACGAGGAGCTGCTCCGCGA 3113

## RESULT 8

US-09-252-991A-198

; Sequence 198, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 198

; LENGTH: 1020

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-198

Query Match 9.6%; Score 36; DB 4; Length 1020;  
Best Local Similarity 47.7%; Pred. No. 1.8;  
Matches 105; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
QY 12 GAGTGAGTACCTGGGCGCAACGAGCGAGGCGGCGGCTGAGTGAAGCTCGAGTT 71  
DB 724 GGGCGCGCTTCGCGCAACGCAAAAAGCCCGCGGATGCGGGGCTTCGAAAGTCGTCT 783  
QY 72 CCGGGTGGCGGGGAGGCGACTGTCCGTGGTCTGAGCGCGCGGCGAGAGCGGGCGCGA 131  
DB 784 GGGCGCGCTCAGGCTGCTTCTGTTCTGCTCCAGGACCTGTCCAGTTCCGCGCAGGC 843  
QY 132 GGGGTGATCGGCTCCCTCGAAGTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGCC 191  
DB 844 GGGCTGATCATCACCTCGGTGATCGGGATCTCCCGGCTCGGCGCTCGATGATGCGCC 903  
QY 192 CCCACCGGCTCCAAAGCTCCAGGCGCTCCCGAGGCAC 231  
DB 904 ACCGGTGTCTGCTGATTTGAAGGGGATCAGCGCGCGC 943

## RESULT 9

US-08-445-515-52/c

; Sequence 52, Application US/08445515

; Patent No. 6043088

; GENERAL INFORMATION:

; APPLICANT: Bookstein, Robert

; APPLICANT: Isaacs, William B.

; TITLE OF INVENTION: A No. 6043088el Prostate/Colon Tumor Suppressor

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/445,515

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-CJ 1607

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

; TELEFAX: (619) 535-8949

; INFORMATION FOR SEQ ID NO: 52:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1342 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-445-515-52

Query Match

Best Local Similarity 9.5%; Score 35.8; DB 3; Length 1342;

Matches 103; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 20 CTTACCTTGGCGCAACGAGCGAGGCGGCGGCGAGTGGAAAGTGGAGTTCCGGGTG 79  
DB 279 CTTCCCGAGCTGATGTCAGAGCAGCAGCAGCAGGAGGAGGAGGAGGAGCTCCCGTG 220  
QY 80 GCGGGGAGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 139  
DB 219 GGCAGGTACCGCAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 160  
QY 140 TCGGCTCCTCGAACTGGGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199  
DB 159 ATCGGGGAGGCGAGTGTCTCTCCACGCGGCTAGCGCTGCGCGCCCGGCTCTACCGC 100  
QY 200 GCTCCAAAGTCCACGGGCTCCCGAGGCACCGG 234  
DB 99 GCACCAACGCTGCTCCCGCGCAGCTCCGCGAGCCG 65

## RESULT 10

US-08-445-515-53/c

; Sequence 53, Application US/08445515

; Patent No. 6043088

; GENERAL INFORMATION:

; APPLICANT: Bookstein, Robert

; APPLICANT: Isaacs, William B.

; TITLE OF INVENTION: A No. 6043088el Prostate/Colon Tumor Suppressor

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92122



QY 73 CGGGTGGCGGGGAGGCGACTGTCTCGTGTGAGCGCGCGAGAGCGGGCGGAG 132  
|||||  
Db 431 CGGGGAGGCGGCGAGCTGAAGGAGGACGCGCTGTGAGCCGAGGAGAGCCCGGCGGAG 372  
|||||  
QY 133 CGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGGCTTAGGGCCCAAGCCC 192  
|||||  
Db 371 TTTTCGGGCGCTTACGCCCGGGGAGGAGCGAGTGGCTCGGGGTGAGGGGAGCCCGGAT 312  
|||||  
QY 193 CCACCGCGCTCCAAAGCTCCGAGGGCGCTCCCGAGGACAGG 234  
|||||  
Db 311 CTTCCGAGAGCCAAAGTCTCTCTCAGCCGCGCGCCGACAG 270  
|||||

## RESULT 13

US-08-600-982-22/c  
; Sequence 22, Application US/08600982  
; Patent No. 6120991  
; GENERAL INFORMATION:  
; APPLICANT: Carter, William G.  
; APPLICANT: Gil, Susanna A.  
; APPLICANT: Ryan, Maureen C.  
; TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness  
; STREET: 1420 Fifth Avenue  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101-8100  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/600,982  
; FILING DATE: 02-SEP-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shelton, Dennis K.  
; REGISTRATION NUMBER: 26,997  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 682-8100  
; TELEFAX: (206) 224-0779  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1994 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; DESCRIPTION: cDNA sequence corresponding to FIGURES

DESCRIPTION: 11A-11C  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
US-08-600-982-22

Query Match 9.4%; Score 35.4; DB 3; Length 1994;  
Best Local Similarity 50.3%; Pred. No. 2.7;  
Matches 87; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 173 TCGCTTAGGGCCCAAGCCCGGCTCCAAAGCTCCAGAGGCGCTCCCGGCGCACC 232  
Db 778 TGGCTGGGTCTCCATGTGCTCATCTGCTCCAGAGGCGCTGCGTGGCACTCAGGCCCT 719  
|||||  
QY 233 GGTGCTCGGCCCTTCCTTCGGTCAGAAAGTCGCCCCCTCGGGGCGAGTTCTCTCCAAAGGG 292  
Db 718 GCAGCTGAGACTTGACCGGCGAGCTGCTGGCCCATGTGCGCCAGTGTCTTCAGAGGG 659  
|||||  
QY 293 TTTCTCTGAAAGAAATCTGAGAGGGCGGCGAGTCTCTTGACCGAGGAATCTCTCTG 345  
|||||

US-08-600-982-22

Query Match 9.4%; Score 35.4; DB 3; Length 1994;  
Best Local Similarity 50.3%; Pred. No. 2.7;  
Matches 87; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 173 TCGCTTAGGGCCCAAGCCCGGCTCCAAAGCTCCAGAGGCGCTCCCGGCGCACC 232  
Db 778 TGGCTGGGTCTCCATGTGCTCATCTGCTCCAGAGGCGCTGCGTGGCACTCAGGCCCT 719  
|||||  
QY 233 GGTGCTCGGCCCTTCCTTCGGTCAGAAAGTCGCCCCCTCGGGGCGAGTTCTCTCCAAAGGG 292  
Db 718 GCAGCTGAGACTTGACCGGCGAGCTGCTGGCCCATGTGCGCCAGTGTCTTCAGAGGG 659  
|||||  
QY 293 TTTCTCTGAAAGAAATCTGAGAGGGCGGCGAGTCTCTTGACCGAGGAATCTCTCTG 345  
|||||

Db 658 TCATCACACAGGTGTGCGCAATCATCATCTCTTCGAGGGTGTGTAICTTG 606  
|||||

## RESULT 14

PCT-US94-10261A-22/c  
; Sequence 22, Application PC/TUS9410261A  
; GENERAL INFORMATION:  
; APPLICANT: Carter, William G.  
; APPLICANT: Gil, Susanna A.  
; APPLICANT: Ryan, Maureen C.  
; TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness  
; STREET: 1420 Fifth Avenue  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101-8100  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/10261A  
; FILING DATE: 02-SEP-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shelton, Dennis K.  
; REGISTRATION NUMBER: 26,997  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 682-8100  
; TELEFAX: (206) 224-0779  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1994 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; DESCRIPTION: cDNA sequence corresponding to FIGURES 11A-11C  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
PCT-US94-10261A-22

Query Match 9.4%; Score 35.4; DB 5; Length 1994;  
Best Local Similarity 50.3%; Pred. No. 2.7;  
Matches 87; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 173 TCGCTTAGGGCCCAAGCCCGGCTCCAAAGCTCCAGAGGCGCTCCCGGCGCACC 232  
Db 778 TGGCTGGGTCTCCATGTGCTCATCTGCTCCAGAGGCGCTGCGTGGCACTCAGGCCCT 719  
|||||  
QY 233 GGTGCTCGGCCCTTCCTTCGGTCAGAAAGTCGCCCCCTCGGGGCGAGTTCTCTCCAAAGGG 292  
Db 718 GCAGCTGAGACTTGACCGGCGAGCTGCTGGCCCATGTGCGCCAGTGTCTTCAGAGGG 659  
|||||  
QY 293 TTTCTCTGAAAGAAATCTGAGAGGGCGGCGAGTCTCTTGACCGAGGAATCTCTCTG 345  
Db 658 TCATCACACAGGTGTGCGCAATCATCATCTCTTCGAGGGTGTGTAICTTG 606  
|||||

US-09-679-279-1/c

; Sequence 1, Application US/09679279  
; Patent No. 6524841  
; GENERAL INFORMATION:  
; APPLICANT: McDaniel, Robert  
; APPLICANT: Volchegursky, Yanina

```

; TITLE OF INVENTION: Recombinant Megalomycin Biosynthetic
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 300622004700
; CURRENT APPLICATION NUMBER: US/09/679,279
; CURRENT FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/158,305
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 60/190,024
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 47981
; TYPE: DNA
; ORGANISM: Micromonospora megalomicea
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(144)
; OTHER INFORMATION: megBVI (megT), TDP-4-keto-6-deoxyglucose-2,3-dehydratase;
; OTHER INFORMATION: SEQ ID NO: 2= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (928)....(2061)
; OTHER INFORMATION: megDVI, TDP-4-keto-6-deoxyglucose 3,4-isomerase;
; OTHER INFORMATION: TDP-4-keto-6-deoxyhexose 3,4-isomerase;
; OTHER INFORMATION: SEQ ID NO: 3= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (2072)....(3382)
; OTHER INFORMATION: megDI, rhodosaminyl transferase (eryCIII homolog);
; OTHER INFORMATION: TDP-megosamine glycosyltransferase;
; OTHER INFORMATION: SEQ ID NO: 4= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (3462)....(4634)
; OTHER INFORMATION: megG (megY), mycarosyl acyltransferase, mycarose O-acyltransferase
; OTHER INFORMATION: SEQ ID NO: 5= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (4651)....(5775)
; OTHER INFORMATION: megDII, deoxysugar transaminase (eryCI, DnrJ homolog);
; OTHER INFORMATION: TDP-3-keto-6-deoxyhexose 3-amino transferase;
; OTHER INFORMATION: SEQ ID NO: 6= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (5822)....(6595)
; OTHER INFORMATION: megDIII, daunosaminyl-N,N-dimethyltransferase (eryCVI homolog);
; OTHER INFORMATION: SEQ ID NO: 7= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (6592)....(7197)
; OTHER INFORMATION: megDIV, TDP-4-keto-6-deoxyglucose 3,5-epimerase (eryBVII, dnmU homolog);
; OTHER INFORMATION: TDP-4-keto-6-deoxyhexose 3,5-epimerase;
; OTHER INFORMATION: SEQ ID NO: 8= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (7220)....(8206)
; OTHER INFORMATION: megDV, TDP-hexose 4-ketoreductase (eryBIV, dnmV homolog);
; OTHER INFORMATION: TDP-4-keto-6-deoxyhexose 4-ketoreductase;
; OTHER INFORMATION: SEQ ID NO: 9= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (8228)....(9220)
; OTHER INFORMATION: megBII-1 (megDVII), TDP-4-keto-L-6-deoxy-hexose 2,3-reductase;
; OTHER INFORMATION: SEQ ID NO: 10= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (9226)....(10479)
; OTHER INFORMATION: megBV, mycarosyl transferase, mycarose glycosyltransferase;
; OTHER INFORMATION: SEQ ID NO: 11= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (10493)....(11424)
; OTHER INFORMATION: megBIV, TDP-hexose 4-ketoreductase;
; OTHER INFORMATION: TDP-4-keto-6-deoxyhexose 4-ketoreductase;
; OTHER INFORMATION: SEQ ID NO: 12= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (12181)....(22821)
; OTHER INFORMATION: megAI, SEQ ID NO: 13= translated amino acid sequence
; NAME/KEY: misc_feature
; LOCATION: (12505)....(13470)
; OTHER INFORMATION: megAI, AT-L
; NAME/KEY: misc_feature
; LOCATION: (37860)....(38120)
;
; LOCATION: (13576)....(13791)
; NAME/KEY: misc_feature
; OTHER INFORMATION: megAI, ACP-I
; LOCATION: (13859)....(15126)
; OTHER INFORMATION: megAI, KS1
; NAME/KEY: misc_feature
; LOCATION: (15427)....(16476)
; OTHER INFORMATION: megAI, AT1
; NAME/KEY: misc_feature
; LOCATION: (17155)....(17694)
; OTHER INFORMATION: megAI, KR1
; NAME/KEY: misc_feature
; LOCATION: (17947)....(18207)
; OTHER INFORMATION: megAI, ACP1
; NAME/KEY: misc_feature
; LOCATION: (18268)....(19548)
; OTHER INFORMATION: megAI, KS2
; NAME/KEY: misc_feature
; LOCATION: (19876)....(20910)
; OTHER INFORMATION: megAI, AT2
; NAME/KEY: misc_feature
; LOCATION: (21517)....(22053)
; OTHER INFORMATION: megAI, KR2
; NAME/KEY: misc_feature
; LOCATION: (22318)....(22575)
; OTHER INFORMATION: megAI, ACP2
; NAME/KEY: CDS
; LOCATION: (22867)....(33555)
; OTHER INFORMATION: megAI; SEQ ID NO: 14= translated amino acid sequence
; NAME/KEY: misc_feature
; LOCATION: (22957)....(24237)
; OTHER INFORMATION: megAI, KS3
; NAME/KEY: misc_feature
; LOCATION: (24544)....(25581)
; OTHER INFORMATION: megAI, AT3
; NAME/KEY: misc_feature
; LOCATION: (26230)....(26733)
; OTHER INFORMATION: megAI, KR3 (inactive)
; NAME/KEY: misc_feature
; LOCATION: (26958)....(27258)
; OTHER INFORMATION: megAI, ACP3
; NAME/KEY: misc_feature
; LOCATION: (27393)....(28590)
; OTHER INFORMATION: megAI, KS4
; NAME/KEY: misc_feature
; LOCATION: (28897)....(29931)
; OTHER INFORMATION: megAI, AT4
; NAME/KEY: misc_feature
; LOCATION: (29553)....(30477)
; OTHER INFORMATION: megAI, DH4
; NAME/KEY: misc_feature
; LOCATION: (31396)....(32244)
; OTHER INFORMATION: megAI, ER4
; NAME/KEY: misc_feature
; LOCATION: (32257)....(32799)
; OTHER INFORMATION: megAI, KR4
; NAME/KEY: misc_feature
; LOCATION: (33052)....(33312)
; OTHER INFORMATION: megAI, ACP4
; NAME/KEY: CDS
; LOCATION: (33666)....(43271)
; OTHER INFORMATION: megAI; SEQ ID NO: 15= translated amino acid sequence
; NAME/KEY: misc_feature
; LOCATION: (33780)....(35027)
; OTHER INFORMATION: megAI, KS5
; NAME/KEY: misc_feature
; LOCATION: (35385)....(36419)
; OTHER INFORMATION: megAI, AT5
; NAME/KEY: misc_feature
; LOCATION: (37068)....(37604)
; OTHER INFORMATION: megAI, KR5
; NAME/KEY: misc_feature
; LOCATION: (37860)....(38120)

```

```
; OTHER INFORMATION: megAIII, ACPS
; NAME/KEY: misc feature
; LOCATION: (38187)...(39470)
; OTHER INFORMATION: megAIII, KS6
; NAME/KEY: misc feature
; LOCATION: (39795)...(40811)
; OTHER INFORMATION: megAIII, AT6
; NAME/KEY: misc feature
; LOCATION: (41406)...(41936)
; OTHER INFORMATION: megAIII, KR6
; NAME/KEY: misc feature
; LOCATION: (42168)...(42425)
; OTHER INFORMATION: megAIII, ACP6
; NAME/KEY: misc feature
; LOCATION: (42585)...(43271)
; OTHER INFORMATION: megAIII, TS
; NAME/KEY: CDS
; LOCATION: (43268)...(44344)
; OTHER INFORMATION: megCII, TDP-4-keto-6-deoxyglucose 3,4-isomerase;
; OTHER INFORMATION: SEQ ID NO: 16= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (44355)...(45623)
; OTHER INFORMATION: megCIII, desosaminyl transferase, desosamine glycosyltransferase;
; OTHER INFORMATION: SEQ ID NO: 17= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (45620)...(46591)
; OTHER INFORMATION: megBII-2(megBII), TDP-4-keto-6-deoxy-L-glucose 2,3 dehydratase,
; OTHER INFORMATION: TDP-4-keto-6-deoxyglucose 2,3 dehydratase;
; OTHER INFORMATION: SEQ ID NO: 18= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (46660)...(47403)
; OTHER INFORMATION: megH, TEII; SEQ ID NO: 19= translated amino acid sequence
; NAME/KEY: CDS
; LOCATION: (47411)...(47980)
; OTHER INFORMATION: megF, C-6 hydroxylase; SEQ ID NO: 20= translated amino acid sequence
US-09-679-279-1
```

```
Query Match          9.4%; Score 35.4; DB 4; Length 47981;
Best Local Similarity 48.3%; Pred. No. 4;
Matches 99; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY   93  CTGTCTGTGTGTAGCCGCGGAGAGCGGCGGAGCGGCTGATCGGCTCCCTCGA 152
Db   13791 CGGGCGCGCAGCAGCGCTCGGTGAGGCGGTGCGGGTCTGATGAGCATGCG 13732

QY   153  ACTGGGGAGTCCAGTGGGTGCTTAGGGCCCAAGCCGCCCGCTCCAAAAGCTC 212
Db   13731 GATGTGAGATCCCGCCCGGTGCGCTGCGAGGCGGTCCGAGCTGCACCGGAGGAC 13672

QY   213  CCAGGGCCTCCCGAGCACCGGTGCTCGGCCCTTCTTCTCGGTGAGAAAGTCCGCCCTGG 272
Db   13671 CGAGTCGAGGCGCAGTTCCCGAAGGTGCGCGGCGTCCGACCTCCGCCCGCGAGGCG 13612

QY   273  GGCAGTTCGTCCCAAGGTTTC 297
Db   13611 GGCCTGTCGCGCGCAGCATTC 13587
```

Search completed: December 5, 2003, 22:59:17  
Job time : 35.0095 secs



QY 181 GCGCCAAAGCCCGCCGCTCCAAAGCTCCAGGGCTCCCGAGGACCGGTGCTCG 240  
Db 181 GCGCCAAAGCCCGCCGCTCCAAAGCTCCAGGGCTCCCGAGGACCGGTGCTCG 240  
QY 241 GCGCTTCCTTCGCTCAGAAAGTCCGCCCCCTGGGGGCGAGTTCGTCACAAAGGGTTTCCTCG 300  
Db 241 GCGCTTCCTTCGCTCAGAAAGTCCGCCCCCTGGGGGCGAGTTCGTCACAAAGGGTTTCCTCG 300  
QY 301 AAGAAATCTGAGAGGGCGAGTCTTGACCGAGGGAATCTCTGTGTAGCCCTTGAAGC 360  
Db 301 AAGAAATCTGAGAGGGCGAGTCTTGACCGAGGGAATCTCTGTGTAGCCCTTGAAGC 360  
QY 361 CGCCAGCCCGCAGAG 375  
Db 361 CGCCAGCCCGCAGAG 375

## RESULT 2

US-10-017-161-1435  
; Sequence 1435, Application US/10017161  
; Publication No. US2003014368A1  
; GENERAL INFORMATION:  
; APPLICANT: SUWA, MAKIKO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: AKIYAMA, YUTAKA  
; APPLICANT: ABURATANI, HIROYUKI  
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
; FILE REFERENCE: 084335/0152  
; CURRENT APPLICATION NUMBER: US/10/017,161  
; CURRENT FILING DATE: 2002-12-18  
; PRIOR APPLICATION NUMBER: JP 2001/246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2430  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1435  
; LENGTH: 922  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(922)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (201)..(722)  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (1)..(26)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (62)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (124)..(125)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (127)..(128)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (131)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (133)..(134)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (138)  
; OTHER INFORMATION: a, t, c, g, unknown or other

; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (142)..(145)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (147)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (151)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (311)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (322)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (326)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (345)..(349)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (351)..(361)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (364)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (366)..(368)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (370)..(380)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (384)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (386)..(388)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (392)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (394)..(396)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (403)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (413)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (415)..(417)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:

```
; NAME/KEY: modified_base
; LOCATION: (422)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (438)..(439)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (441)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (443)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (445)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (465)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (469)..(471)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (487)..(488)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (490)..(491)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (495)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (497)..(498)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (501)..(505)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (519)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (529)..(530)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (540)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (548)..(549)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (564)..(568)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (571)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
```

```
; LOCATION: (573)
Query Match 11.0%; Score 41.2; DB 12; Length 922;
Best Local Similarity 40.9%; Pred. No. 0.047;
Matches 70; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 12 GAGCTGAGCTTACCTTGGCGCAACGACGAGGCGAGCGCGAGTGAAGCTGAGATT 71
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
363 GNGNNGNNNNNNNNNNNGGNGNNGGNGGCGGCGGCGGCGGCGGCGGCGGCGG 422

QY 72 CCGGGGTGGCGGAGGAGCTGCTCCGTGCTGAGCGCCGCGGAGCGCGCGGA 131
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
423 GCGGGCGGCGGCGGCGGNGNGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 482

QY 132 GCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTCCGCTTAGGG 182
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
483 GCGGNGNGGNGGNGNNGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 533
```

## RESULT 3

```
US-09-771-161A-58
; Sequence 58, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-771-161A-58
```

```
Query Match 10.9%; Score 41; DB 10; Length 1097;
Best Local Similarity 52.0%; Pred. No. 0.053;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
```

```
QY 1 CCCTCCCTTCCGAGCTGAGCTTACCTTGGCGCGCAACGAGCGAGGCGGCGCGAGTGG 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23 CCCCCGGCTCGGGGCTGTGAGCGCTCGCGGCGCGGCGGCGGCGGCGGCGG 82

QY 61 AAGCTGAGTTCGGGGTGGCGGGGAGCGGAGCTGTCCGTGCTGAGCGCGCGGAG 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 CCGACGCTCTCTTCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 142

QY 121 GCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTGCCT 177
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
143 GCGGGCGGAGTTCCTCCCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 199
```

## RESULT 4

```
US-09-771-161A-59
; Sequence 59, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
```

;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: 135619  
;; PRIOR FILING DATE: 2000-04-12  
;; NUMBER OF SEQ ID NOS: 273  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 59  
;; LENGTH: 3382  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: -  
;; LOCATION: (1)..(3382)  
;; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'  
US-09-771-161A-59

Query Match 10.9%; Score 41; DB 10; Length 3382;  
Best Local Similarity 52.0%; Pred. No. 0.05;  
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
QY 1 CCTCTCCCTCCGAGCTGAGCTTACCTGGGCGCAACGAGGAGGAGGCGGCGCGAGTGG 60  
DB 23 CCCCCGGCTCGGGCTGTGAGCGGCTCGGGGCGGGGGTGGCGGCGGTGCGGCGGGCG 82  
QY 61 AAGCTGAGATTCCGGGGTGGGGGAGCGCACTGTCCGTGTGTGTGAGCGCGCGGAGGA 120  
DB 83 CGGACGCTCTCTTCGGCGGCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 142  
QY 121 GCGGGCGGAGCGGCTGATCGGCTCCCTCGACTGGGAGAGTCCAGTGGGGTTCGCT 177  
DB 143 GCGGGCGGAGGTTGCTCTCCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 199

## RESULT 5

US-09-291-417-102  
;; Sequence 102, Application US/09291417A  
;; Publication No. US20030050230A1  
;; GENERAL INFORMATION:  
;; APPLICANT: PLOWMAN, GREGORY  
;; APPLICANT: MARTINEZ, RICARDO  
;; APPLICANT: WHYTE, DAVID  
;; TITLE OF INVENTION: STR20-RELATED PROTEIN KINASES  
;; FILE REFERENCE: 240/300  
;; CURRENT APPLICATION NUMBER: US/09/291,417A  
;; CURRENT FILING DATE: 1999-04-13  
;; EARLIER APPLICATION NUMBER: US 60/081,784  
;; EARLIER FILING DATE: 1998-04-14  
;; NUMBER OF SEQ ID NOS: 147  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 102  
;; TYPE: DNA  
;; ORGANISM: Full Length Mammalian (Human) PAK5

US-09-291-417-102  
Query Match 10.7%; Score 40.2; DB 11; Length 2806;  
Best Local Similarity 49.8%; Pred. No. 0.083;  
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 23 ACCCTGGGCGCAACGAGCGAGGCGGCGAGTGGAGCTGAGTTCGGGGTGGGC 82  
DB 567 ACCACGGCCAGAGGGGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 626  
QY 83 GGGGAGGCGACTGTCCGTGTGTCTGAGCGCGGAGAGCGGCGGCGGAGCGGCTGTATCG 142  
DB 627 GAGCGGGTGGGCGGAGTGGTACAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 686  
QY 143 GCTCCCTCGACTCGAGGAGTCCAGTGGGGTGTCTTAGGGCCCAAGCCCCACCGGCT 202  
DB 687 TCCAGGAGGGGCTCAGGGGGTCCCGAGGAGTCTCCCGGAGCAACAGCCCTCTCCGGG 746  
QY 203 CCAAAAGCTCCAGGGGCTCCCGAG 227  
DB 747 CTGTATGTGGGACACCCCGGCTG 771

## RESULT 6

US-10-134-102-3  
;; Sequence 3, Application US/10134102  
;; Publication No. US20030186254A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Melnick, Michael B.  
;; APPLICANT: Moritz, Albrecht  
;; APPLICANT: Comb, Michael J.  
;; TITLE OF INVENTION: Regulation of HIV-Tat and Nef by the Pak4 kinase and its  
;; FILE REFERENCE: CST-176 CIP  
;; CURRENT APPLICATION NUMBER: US/10/134,102  
;; PRIOR FILING DATE: 2002-04-29  
;; PRIOR APPLICATION NUMBER: 09/750,457  
;; PRIOR FILING DATE: 2000-12-28  
;; PRIOR APPLICATION NUMBER: 60/173,939  
;; PRIOR FILING DATE: 1999-12-30  
;; NUMBER OF SEQ ID NOS: 10  
;; SOFTWARE: PatentIn ver. 2.0  
;; SEQ ID NO 3  
;; LENGTH: 2838  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (211)..(1986)  
US-10-134-102-3

Query Match 10.7%; Score 40.2; DB 12; Length 2838;  
Best Local Similarity 49.8%; Pred. No. 0.083;  
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 23 ACCCTGGGCGCAACGAGCGAGGCGGCGGAGTGGAGCTGAGTTCGGGGTGGGC 82  
DB 577 ACCACGGCCAGAGGGGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 636  
QY 83 GGGGAGGCGACTGTCCGTGTGTCTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 142  
DB 637 GAGCGGGTGGGCGGAGTGGTGTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGTCT 696  
QY 143 GCTCCCTCGACTCGGAGGAGTGGAGTGGGTGTCTTAGGGCCCAAGCCCCACCGGCT 202  
DB 697 TCCAGGAGGAGGCTCAGGGGTCCCGAGGAGTCTCCCGGAGCAACAGCCCTCTCCGGG 756  
QY 203 CCAAAAGCTCCAGGGGCTCCCGAG 227  
DB 757 CTGTATGTGGGACACCCCGGCTG 781

## RESULT 7

US-09-960-352-2238  
;; Sequence 2238, Application US/09960352  
;; Patent No. US20020137139A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Warren, Wesley C.  
;; APPLICANT: Tao, Mengbing  
;; APPLICANT: Byatt, John C.  
;; APPLICANT: Mathialagan, Nagappan  
;; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
;; FILE REFERENCE: 15511, 006/37-21(10298)C  
;; CURRENT APPLICATION NUMBER: US/09/960,352  
;; CURRENT FILING DATE: 2001-09-24  
;; NUMBER OF SEQ ID NOS: 15112  
;; SEQ ID NO 2238  
;; LENGTH: 392  
;; TYPE: DNA  
;; ORGANISM: Bos taurus  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (270)

US-09-960-352-2238  
Query Match 10.7%; Score 40.2; DB 11; Length 2806;  
Best Local Similarity 49.8%; Pred. No. 0.083;  
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 23 ACCCTGGGCGCAACGAGCGAGGCGGCGAGTGGAGCTGAGTTCGGGGTGGGC 82  
DB 567 ACCACGGCCAGAGGGGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 626  
QY 83 GGGGAGGCGACTGTCCGTGTGTCTGAGCGCGGAGAGCGGCGGCGGAGCGGCTGTATCG 142  
DB 627 GAGCGGGTGGGCGGAGTGGTACAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 686  
QY 143 GCTCCCTCGACTCGAGGAGTCCAGTGGGGTGTCTTAGGGCCCAAGCCCCACCGGCT 202  
DB 687 TCCAGGAGGGGCTCAGGGGGTCCCGAGGAGTCTCCCGGAGCAACAGCCCTCTCCGGG 746  
QY 203 CCAAAAGCTCCAGGGGCTCCCGAG 227  
DB 747 CTGTATGTGGGACACCCCGGCTG 771

OTHER INFORMATION: unsure at all n locations  
OTHER INFORMATION: Clone ID: 10-LIB3058-057-Q1-K1-C5  
US-09-960-352-2238

Query Match 10.6%; Score 39.8; DB 10; Length 392;  
Best Local Similarity 52.4%; Pred. No. 0.12; Mismatches 0; Gaps 0;  
Matches 86; Conservative 0; Indels 78; Gaps 0;

QY 41 CGAGGACAGGGCGGAGTGGAGTTCGGGGTGGGGAGGAGCGACTGTCGGT 100  
DB 186 CGAGGAGAGGAGCAGGAGGAGCGGTGCGGGTCCGGTGGCGGCCCAAGTCCG 245  
QY 101 GGTGTCGAGCGCGGAGAGCGGGCGGAGCGGCTGATCGCTCCCTCGAACTGGGGA 160  
DB 246 AGGCTTACCGCGCGCGCGCGGCGCGCAGGGCCCTGGCGGAGGCGCCCGCGGGA 305  
QY 161 GTTCCAGTGGGTGCGTTAGGGCCCAAGCCCGCCAGCCCGGCTCC 204  
DB 306 GTCCACGAGGCGCGGTTCCGCGCGCGGACCCCGCGCGGACC 349

## RESULT 8

US-10-029-386-22980/c

Sequence 22980, Application US/10029386

Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 22980

LENGTH: 587

TYPE: DNA

FEATURE:

ORGANISM: Homo sapiens

OTHER INFORMATION: MAP TO AL035703.17

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7

OTHER INFORMATION: SWISSPROT HIT: P19332, EVALUE 2.70e+00

OTHER INFORMATION: EST HUMAN HIT: AL528782.1, EVALUE 2.00e-38

OTHER INFORMATION: NT HIT: AF210249.1, EVALUE 2.00e-03

US-10-029-386-22980

Query Match 10.4%; Score 39; DB 12; Length 587;

Best Local Similarity 48.7%; Pred. No. 0.19;

Matches 135; Conservative 0; Mismatches 140; Indels 2; Gaps 1;

QY 26 CTGGGCGCAACAGCAGGAGCGGCGGAGTGAAGTTCGGGGTGGCGGG 85

DB 290 CCGGCCCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 231

QY 86 GAGCGACTTCGTTGAGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGG 145

DB 230 GAGCGCGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 171

QY 146 CCCT--CGAATGGGAGTCCAGTGGGGTTCGTTAGGCGCCAAAGCCCGCGGCTC 203

DB 170 TCCCTCCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 111

QY 204 CAAAGTCCAGCGGCTCCCGAGGCAACCGGCTTCGCGCCCTTCCTTCGAGAGTC 263

DB 110 CCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 51

QY 264 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300

DB 50 CACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 14

## RESULT 9

US-09-974-300-836

Sequence 836, Application US/09974300

Patent No. US20020146721A1

GENERAL INFORMATION:

APPLICANT: Berka, Randy M.

APPLICANT: Clausen, Ib Groth

TITLE OF INVENTION: Methods For Monitoring Multiple Gene

FILE REFERENCE: 10085.500-US

CURRENT APPLICATION NUMBER: US/09/974,300

CURRENT FILING DATE: 2001-10-05

PRIOR APPLICATION NUMBER: 09/680,598

PRIOR FILING DATE: 2000-10-06

PRIOR APPLICATION NUMBER: 60/279,526

PRIOR FILING DATE: 2001-03-27

NUMBER OF SEQ ID NOS: 8481

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 836

LENGTH: 936

TYPE: DNA

ORGANISM: Bacillus licheniformis

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)---(936)

OTHER INFORMATION: n = A,T,C or G

US-09-974-300-836

Query Match 10.4%; Score 39; DB 10; Length 936;

Best Local Similarity 55.6%; Pred. No. 0.19;

Matches 75; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 38 GAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 97

DB 405 GAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 464

QY 98 CGTGTGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 157

DB 465 GGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 524

QY 158 GAGGTTCAGTGGG 172

DB 525 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 539

## RESULT 10

US-10-029-386-24925/c

Sequence 24925, Application US/10029386

Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 24925

LENGTH: 550

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL138846.2

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.65

OTHER INFORMATION: SWISSPROT HIT: Q35052, EVALUE 4.00e+00

OTHER INFORMATION: NT HIT: G113650682, EVALUE 8.20e+00

OTHER INFORMATION: EST\_HUMAN HIT: BF745459.1, EVALUE 1.00e-103

## US-10-029-386-24925

Query Match 10.3%; Score 38.6; DB 12; Length 550;  
Best Local Similarity 52.9%; Pred. No. 0.25;  
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
QY 27 TGGCGCGCAACGAGCGAGGCGGCGGAGTGGAGCTTCGGGGTGGCGGGG 86  
Db 305 TAGACGCGCAGGCTTCGACGCGGGCGGGTCCGGCTCCGGGTGCGGCCCGAGGGAG 246  
QY 87 AGCGCACTGTCCTGTGTGCTGAGCGCGCGGAGAGCGGCGGCGGAGCGGCTGATCGGCTC 146  
Db 245 CCGCCAGGTCCTCGGTGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCC 186  
QY 147 CCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAGGGC 183  
Db 185 GGGGCCCTGCGACGCGGCGCGGGGGCCCCCTCTGGGC 149

## RESULT 11

US-10-120-988-142/c  
; Sequence 142, Application US/10120988  
; Publication No. US20030219745A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030219745A1e1 Nucleic Acids and  
; FILE REFERENCE: 802CON  
; CURRENT APPLICATION NUMBER: US/10/120,988  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 09/774,528  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: pt.FL\_Genes Version 2.0  
; SEQ ID NO 142  
; LENGTH: 2310  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1374)  
US-10-120-988-142

Query Match 10.3%; Score 38.6; DB 12; Length 2310;  
Best Local Similarity 52.9%; Pred. No. 0.23;  
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
QY 27 TGGCGCGCAACGAGCGAGGCGGCGGAGTGGAGCTTCGGGGTGGCGGGG 86  
Db 305 TAGACGCGCAGGCTTCGACGCGGGCGGGTCCGGCTCCGGGTGCGGCCCGAGGGAG 246  
QY 87 AGCGCACTGTCCTGTGTGCTGAGCGCGGCGGAGCGGCGGCGGAGCGGCTGATCGGCTC 146  
Db 245 CCGCCAGGTCCTCGGTGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCC 186  
QY 147 CCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAGGGC 183  
Db 185 GGGGCCCTGCGACGCGGCGCGGGGGCCCCCTCTGGGC 149

## RESULT 12

US-10-017-161-1857  
; Sequence 1857, Application US/10017161  
; Publication No. US20030143668A1  
; GENERAL INFORMATION:  
; APPLICANT: SUWA, MAKIHO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI  
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
; FILE REFERENCE: 084335/0152  
; CURRENT APPLICATION NUMBER: US/10/017,161  
; CURRENT FILING DATE: 2002-12-18  
; PRIOR APPLICATION NUMBER: JP 2001/246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2430  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1857  
; LENGTH: 3163  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(3163)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (201)..(2963)  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (556)..(655)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (658)..(669)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (671)..(697)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (704)..(707)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (709)..(736)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (739)..(743)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (745)..(747)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (750)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (752)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (754)..(756)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (761)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (763)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (818)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified\_base

LOCATION: (843)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (892)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (897)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (920)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (941)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (993)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (995)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1066)..(1067)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1082)..(1084)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1105)..(1107)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1130)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1147)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1162)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1202)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1233)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1247)..(1248)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1261)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1485)	a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1500)	a, t, c, g, unknown or other

; APPLICANT: MOORE, Jeffrey  
 ; APPLICANT: BURMAN, Ed T.  
 ; APPLICANT: BRADLEY, John  
 ; APPLICANT: DESILVA, Tamara  
 ; APPLICANT: HARRIS, Sandra  
 ; APPLICANT: KOMARNILSKY, Svetlana  
 ; APPLICANT: MENDILLO, Marc  
 ; APPLICANT: MOORE, Daniel  
 ; APPLICANT: MCCOY, Melissa  
 ; APPLICANT: SANDERSON, Karen  
 ; APPLICANT: HAO, Tariq  
 ; APPLICANT: ZHU, Shuhao  
 ; APPLICANT: LONG, Fan  
 ; APPLICANT: DAVIDOV, Eugene  
 ; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE  
 ; FILE REFERENCE: 0342/1G548-US2  
 ; CURRENT APPLICATION NUMBER: US/09/893,519A  
 ; CURRENT FILING DATE: 2001-06-28  
 ; PRIOR APPLICATION NUMBER: US 60/215,164  
 ; PRIOR FILING DATE: 2000-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/224,457  
 ; PRIOR FILING DATE: 2000-08-10  
 ; NUMBER OF SEQ ID NOS: 146  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 87  
 ; LENGTH: 2307  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; PUBLICATION INFORMATION:  
 ; DATABASE ACCESSION NUMBER: Human Genbank/Y11354.1  
 ; DATABASE ENTRY DATE: 1997-06-25  
 ; RELEVANT RESIDUES: (1)..(2307)  
 US-09-893-519A-87

Query Match 10.1%; Score 38; DB 11; Length 2307;  
 Best Local Similarity 56.3%; Pred. No. 0.34;  
 Matches 71; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
 QY 37 CGACGAGGAGGCGGCGGAGTGAAGCTGGAGTTCCGGGGTGGGGGAGGCGGACTGT 96  
 DB 895 CGGGGGCGGACGCGCGGGCGCGGCTCGGGGGTCCGGGGGTGGCGGGCGCGGG 836  
 QY 97 CCGTGGTGTCTAGCGCGCGGAGAGCGCGGCGGCGGCTGATCGGCTCCCTCGAATCG 156  
 DB 835 CCAGGCTGGCGGCGGCGGCGGCTGGCGGGCGGCGGCGGCGGCGGCGGCGGCGG 776  
 QY 157 GGGAGG 162  
 DB 775 GGGCGG 770

RESULT 14  
 US-09-899-634A-1  
 ; Sequence 1, Application US/09899634A  
 ; Patent No. US20020059654A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thomas Buehler, Reto Andreas Gadiant, Reinhard Korn, Rao Movva  
 ; TITLE OF INVENTION: PCR and its uses  
 ; FILE REFERENCE: 4-31499A  
 ; CURRENT APPLICATION NUMBER: US/09/899,634A  
 ; CURRENT FILING DATE: 2001-07-05  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 4286  
 ; TYPE: DNA  
 ; ORGANISM: porcine  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (3229)..(4014)  
 US-09-899-634A-1

Query Match 10.1%; Score 37.8; DB 9; Length 4286;  
 Best Local Similarity 56.3%; Pred. No. 0.34;  
 Matches 71; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Best Local Similarity 50.3%; Pred. No. 0.37;  
 Matches 93; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
 QY 38 GAGCGAGGAGGCGGCGGAGTGAAGCTGGAGTTCCGGGGTGGGGGAGGCGGACTGTC 97  
 DB 2685 GAGCAGGCGCGGCTTCGGGTGGGGGCTCCGTGCGGGGAGGCTCGCGGCTCGCGG 2744  
 QY 98 CGTGGTCTGAGCGCGCGGCGGAGCGGCGCGGAGGCGGCTGATCGGCTCCCTCGAATCG 157  
 DB 2745 CCGGGCGGCGGCTGGCGGAGGCTGGGGGCTCCGGGGGCGGCGGCGGCTCGCGGCGG 2804  
 QY 158 GGAGTCCAGTGGGGTGGCTTAGGGCCCAAGCCCGGCTCCAAAAGCTCCCGAGG 217  
 DB 2805 GAGGCTCGGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCTGTTCGAGGCGGCGG 2864  
 QY 218 GCCTC 222  
 DB 2865 GCGGC 2869

RESULT 15  
 US-10-161-403-70  
 ; Sequence 70, Application US/10161403  
 ; Publication No. US20030119104A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Perkins, Edward  
 ; APPLICANT: Lindenberg, Michael  
 ; APPLICANT: Greene, Amy  
 ; APPLICANT: Leung, Josephine  
 ; APPLICANT: Fleming, Elena  
 ; APPLICANT: Stewart, Sandra  
 ; APPLICANT: Shellard, Jean  
 ; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
 ; FILE REFERENCE: 24601-420  
 ; CURRENT APPLICATION NUMBER: US/10/161,403  
 ; CURRENT FILING DATE: 2002-05-30  
 ; PRIOR APPLICATION NUMBER: 60/294,758  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: 60/366,891  
 ; PRIOR FILING DATE: 2002-03-21  
 ; NUMBER OF SEQ ID NOS: 129  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 70  
 ; LENGTH: 4778  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: pcx plasmid  
 US-10-161-403-70

Query Match 10.1%; Score 37.8; DB 14; Length 4778;  
 Best Local Similarity 50.3%; Pred. No. 0.37;  
 Matches 93; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
 QY 38 GAGCGAGGAGGCGGCGGAGTGAAGCTGGAGTTCCGGGGTGGGGGAGGCGGACTGTC 97  
 DB 1156 GAGCAGGCGCGGCTTCGGGTGGGGGCTCCGTGCGGGGAGGCTCGCGGCTCGCGG 1215  
 QY 98 CGTGGTCTGAGCGCGCGGCGGAGCGGCGGAGCGGCTGATCGGCTCCCTCGAATCG 157  
 DB 1216 CCGGGCGGCGGCTGGCGGAGGCTGGGGGCTCCGGGGGCGGCGGCGGCTCGCGG 1275  
 QY 158 GGAGTCCAGTGGGGTGGCTTAGGGCCCAAGCCCGGCTCCAAAAGCTCCCGAGG 217  
 DB 1276 GAGGCTCGGGGAGGCGGCGGCGGCGGCGGCGGCGGCTGTTCGAGGCGGCGG 1335  
 QY 218 GCCTC 222  
 DB 1336 GCGGC 1340

Search completed: December 5, 2003, 23:17:02

/ Mon Dec 8 11:25:47 2003

us-09-936-680-4\_copy\_1\_375.rnpb

Page 9

Job time : 157.126 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:16:11 ; Search time 993.523 Seconds  
(without alignment)  
9799.969 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_1024\_1261  
Perfect score: 238  
Sequence: 1 acacagagcagtgatgacatg.....tgcacgaactgagaagcgg 238

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hgt: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vt: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vt: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrt: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	238	100.0	1261	6	AX039100	AX039100 Sequence
2	236	100.0	1261	9	HSA243396	AJ243396 Homo sapi
3	236	99.2	4052	9	HSM801563	AL136589 Homo sapi
4	236	99.2	5306	9	AB032984	AB032984 Homo sapi
5	216	90.8	127347	2	AP000682	AP000682 Homo sapi
6	216	90.8	144833	2	AC063921	AC063921 Homo sapi
7	216	90.8	149800	2	AC021981	AC021981 Homo sapi
8	216	90.8	172546	2	AC024604	AC024604 Homo sapi
9	216	90.8	178169	9	AP002765	AP002765 Homo sapi
10	216	90.8	178431	2	AP002749	AP002749 Homo sapi
11	216	90.8	181471	9	AC069539	AC069539 Homo sapi
12	203.6	85.5	3296	9	AB097521	AB097521 Macaca fa
13	108	45.4	66980	2	AC013796	AC013796 Homo sapi
14	47.4	19.9	189486	2	AC114854	AC114854 Rattus no
15	44	18.5	167342	2	AP006477	AP006477 Homo sapi
16	44	18.5	182917	9	AL139383	AL139383 Human DNA
17	43.8	18.4	181290	2	EX005286	EX005286 Danio rer
18	43.2	18.2	74549	9	HSJ890015	AL049540 Human DNA
19	43.2	18.2	139084	5	AL935332	AL935332 Zebrafish
20	43.2	18.2	192360	2	AC120004	AC120004 Mus muscu
21	43.2	18.2	196687	10	AC122466	AC122466 Mus muscu
22	42.6	17.9	206768	2	AC125951	AC125951 Rattus no
23	42.6	17.9	228381	2	AC103262	AC103262 Rattus no
24	42.6	17.9	240645	2	AC126633	AC126633 Rattus no
25	42.4	17.8	61434	9	AC007366	AC007366 Homo sapi
26	42.4	17.8	93599	2	AC084728	AC084728 Homo sapi
27	42.4	17.8	118674	2	AC023063	AC023063 Homo sapi
28	42.4	17.8	225490	2	AC140213	AC140213 Mus muscu
29	42.4	17.8	239574	2	AC132010	AC132010 Homo sapi
30	42.2	17.7	129502	9	HSBJ726C3	AL121756 Human DNA
31	41.8	17.6	159064	9	AC138512	AC138512 Homo sapi
32	41.6	17.5	205610	9	AC134312	AC134312 Homo sapi
33	41.6	17.5	196823	2	EX323553	EX323553 Danio rer
34	41.6	17.5	196882	2	EX323552	EX323552 Danio rer
35	41.6	17.5	247983	2	EX323588	EX323588 Danio rer
36	41.4	17.4	99958	9	AL606510	AL606510 Human DNA
37	41.4	17.4	226950	10	AL671011	AL671011 Mouse DNA
38	41.2	17.3	151804	2	AC008161	AC008161 Mus muscu
39	41.2	17.3	180745	9	AL359183	AL359183 Human DNA
40	41.2	17.3	205379	10	AL591174	AL591174 Mouse DNA
41	41	17.2	115398	9	AL672311	AL672311 Human DNA
42	41	17.2	155192	2	EX510929	EX510929 Danio rer
43	41	17.2	227750	5	AL844563	AL844563 Zebrafish
44	41	17.2	229552	2	AC128405	AC128405 Rattus no
45	41	17.2	243892	2	AC097933	AC097933 Rattus no

# ALIGNMENTS

RESULT 1  
AX039100  
LOCUS AX039100 1261 bp DNA linear PAT 18-NOV-2000  
DEFINITION Sequence 4 from Patent WO0063367.  
ACCESSION AX039100  
VERSION AX039100.1 GI:11229276  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Cox P., Dixon A., Jackson A. and Morgan K.  
TITLE A novel family of beta sub-unit proteins from a voltage-gated sodi  
um channel, nucleic acids encoding them and therapeutic or

diagnostic uses there of  
 Patent: WO 0063367-A 4 26-OCT-2000;  
 WARNER-LAMBERT COMPANY (US); Cambridge University Technical  
 Services Limited (GB)

FEATURES  
 source  
 1. .1261  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /db\_xref="taxon:9606"  
 /tissue\_type="brain"  
 1. .1261

BASE COUNT 263 a 366 c 372 g 260 t

ORIGIN

Query Match 100.0%; Score 238; DB 6; Length 1261;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-59;  
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGACTGGACATCCCATGTTTC 60  
 |||||  
 Db 1024 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGACTGGACATCCCATGTTTC 1083

QY 61 AGCAATGTCAATGGCATACGAGGCGGCCCAAGGCCCATCGCTTCCCTTCATGCATC 120  
 |||||  
 Db 1084 AGCAATGTCAATGGCATACGAGGCGGCCCAAGGCCCATCGCTTCCCTTCATGCATC 1143

QY 121 CATGTCTCTTCATTCATTCATCCATACATCCATCCCTCTGAGCTTTCACCTCTGAC 180  
 |||||  
 Db 1144 CATGTCTCTTCATTCATTCATCCATACATCCATCCCTCTGAGCTTTCACCTCTGAC 1203

QY 181 TCCTTAACCTCATCAGACCTCTACGACCATACAGACTCTGCAGAACTGAGAGCGCG 238  
 |||||  
 Db 1204 TCCTTAACCTCATCAGACCTCTACGACCATACAGACTCTGCAGAACTGAGAGCGCG 1261

RESULT 2  
 HSA243396 1261 bp mRNA linear PRI 13-MAR-2000  
 LOCUS  
 DEFINITION Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit  
 (scn3b gene).  
 ACCESSION AJ243396.2 GI:7242612  
 VERSION  
 KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K.,  
 Pincock, R.D., Hughes, J., Richardson, P.J., Mizuguchi, K. and  
 Jackson, A.P.  
 TITLE beta 3: an additional auxiliary subunit of the voltage-sensitive  
 sodium channel that modulates channel gating with distinct kinetics  
 Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)  
 MEDLINE 20160948  
 PUBMED 10688874  
 REFERENCE 2  
 Morgan, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JUN-1999) Morgan K., Biochemistry, University of  
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
 REVISION 3 (bases 1 to 1261)  
 REFERENCE 3 (bases 1 to 1261)  
 Morgan, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-MAR-2000) Morgan K., Biochemistry, University of  
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM  
 COMMENT On Mar 14, 2000 this sequence version replaced gi:7160974.  
 FEATURES  
 source  
 1. .1261  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="brain"  
 1. .1261

/gene="scn3b"  
 <1. .375  
 /gene="scn3b"  
 376. .1024  
 /gene="scn3b"  
 /function="modulates channel gating kinetics"  
 /codon\_start=1  
 /evidence="experimental"  
 /product="voltage-gated sodium channel beta-3 subunit"  
 /protein\_id="CA76825.1"  
 /db\_xref="GI:7160975"  
 /translation="MPAPNRLFPPLASLVLIVYVSCFPCVCEVPSSEAVQCNPKLR  
 CISCMPKREVEATTVVEYFPEGRGKDFLIYRKHQEVESFPQGRLOWNGSKDLQD  
 VSTVLNVLTNDISGLYTCNVSEFEFAHRPFVTRILPLRVTEAGEDFTSVVSEI  
 MMYILLVFLTLWLLIEMIYCYRKVKSAEEAQAENASDYLAIPSENKENSAPVPE"  
 1025. .>1261  
 /gene="scn3b"  
 263 a 366 c 372 g 260 t

BASE COUNT 263 a 366 c 372 g 260 t

ORIGIN

Query Match 100.0%; Score 238; DB 9; Length 1261;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-59;  
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGACTGGACATCCCATGTTTC 60  
 |||||  
 Db 1024 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGACTGGACATCCCATGTTTC 1083

QY 61 AGCAATGTCAATGGCATACGAGGCGGCCCAAGGCCCATCGCTTCCCTTCATGCATC 120  
 |||||  
 Db 1084 AGCAATGTCAATGGCATACGAGGCGGCCCAAGGCCCATCGCTTCCCTTCATGCATC 1143

QY 121 CATGTCTCTTCATTCATTCATCCATACATCCATCCCTCTGAGCTTTCACCTCTGAC 180  
 |||||  
 Db 1144 CATGTCTCTTCATTCATTCATCCATACATCCATCCCTCTGAGCTTTCACCTCTGAC 1203

QY 181 TCCTTAACCTCATCAGACCTCTACGACCATACAGACTCTGCAGAACTGAGAGCGCG 238  
 |||||  
 Db 1204 TCCTTAACCTCATCAGACCTCTACGACCATACAGACTCTGCAGAACTGAGAGCGCG 1261

RESULT 3  
 HSM801563 4052 bp mRNA linear PRI 20-MAR-2002  
 LOCUS  
 DEFINITION Homo sapiens mRNA; cDNA DKF2p761F182 (from clone DKF2p761F182);  
 complete cds.  
 ACCESSION AL136589  
 VERSION  
 KEYWORDS AL136589.1 GI:13276580  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 4052)  
 Ordenwaelder, B., Obermaier, B., Mewes, H.W., Weil, B. and Wiemann, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152  
 Martinsried, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by MediGenomix (Martinsried/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. This clone  
 (DKF2p761F182) is available at the RZPD in Berlin. Please contact  
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at http://mips.gsf.de/proj/cDNA/.

FEATURES  
 source  
 1. .4052  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKF2p761F182"

```

/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..4052
/gene="DKF2p761F182"
804..1451
CDS
/gene="DKF2p761F182"
/translation="NPAENRFLPGLASLVLIYVWSVCFVPCVVESETEAVQGNPKLR
CISMKREVEATTVNPFYRPEGKDFLIYVRNGHGVESFPQRLWNGSKLOD
VSTIVLVNLTNDGLYTCNVREFEFAHRPFVKTRLIPLRVTEAGEDFTSVVSEI
MMVILLVFLTLMLLIEMICYRKVSKAEAAQENASDYLAIPEKNSAVPVEE"
polyA_site
4035
/gene="DKF2p761F182"
BASE COUNT 994 a 1033 c 1036 g 989 t
ORIGIN
Query Match 99.2%; Score 236; DB 9; Length 4052;
Best Local Similarity 100.0%; Pred. No. 9.8e-59;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGGACATCCCATGTC 60
Db 1452 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGGACATCCCATGTC 1511
QY 61 AGCAATGTCAATGGCATCAGGAGGGGCCGCCAAGGGGCCCATCGCTTCCCTTCATGCATC 120
Db 1512 AGCAATGTCAATGGCATCAGGAGGGGCCGCCAAGGGGCCCATCGCTTCCCTTCATGCATC 1571
QY 121 CATGTGTCGTTTCATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 180
Db 1572 CATGTGTCGTTTCATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1631
QY 181 TCCTTAACCTCAGACCTTACGACCATTAAGACTTCGCGAGAACTGAGAAGCC 236
Db 1632 TCCTTAACCTCAGACCTTACGACCATTAAGACTTCGCGAGAACTGAGAAGCC 1687

RESULT 4
AB032984
LOCUS
DEFINITION
Homo sapiens mRNA for KIAA1158 protein, partial cds.
ACCESSION
AB032984
VERSION
AB032984.1 GI:6330135
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
Hiroseawa, M., Nagase, T., Ishikawa, K., Kikuno, R., Nomura, N. and
Ohara, O.
Characterization of cDNA clones selected by the Genemark analysis
from size-fractionated cDNA libraries from human brain
DNA Res. 6 (5), 329-336 (1999)
20039618
MEDLINE
10574461
PUBMED
REFERENCE
2 (bases 1 to 5306)
Ohara, O., Nagase, T. and Kikuno, R.
Direct Submission
AUTHORS
Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
282-0812, Japan (E-mail:cdna@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
Location/Qualifiers
1..5306

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bj00081"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
/dev_stage="adult"
1..5306
/gene="KIAA1158"
1..695
CDS
/gene="KIAA1158"
/translation="Start codon is not identified."
/codon_start=3
/product="KIAA1158 protein"
/protein_id="BAA86472.1"
/db_xref="GI:6330136"
/translation="RSWGGSSLGSRQPKMFAFNLFLASLVLIYVWSVCFVPCVVESETEAVQGNPKLRCTSCMKREVEATTVNPFYRPEGKDFLIYVRNGHGVESFPQRLWNGSKLODVSITIVLVNLTNDGLYTCNVREFEFAHRPFVKTRLIPLRVTEAGEDFTSVVSEIMMYILLVFLTLMLLIEMICYRKVSKAEAAQENASDYLAIPEKNSAVPVEE"
BASE COUNT 1375 a 1311 c 1111 g 1509 t
ORIGIN
Query Match 99.2%; Score 236; DB 9; Length 5306;
Best Local Similarity 100.0%; Pred. No. 9.6e-59;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGGACATCCCATGTC 60
Db 696 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGGACATCCCATGTC 755
QY 61 AGCAATGTCAATGGCATCAGGAGGGGCCGCCAAGGGGCCCATCGCTTCCCTTCATGCATC 120
Db 756 AGCAATGTCAATGGCATCAGGAGGGGCCGCCAAGGGGCCCATCGCTTCCCTTCATGCATC 815
QY 121 CATGTGTCGTTTCATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 180
Db 816 CATGTGTCGTTTCATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 875
QY 181 TCCTTAACCTCAGACCTTACGACCATTAAGACTTCGCGAGAACTGAGAAGCC 236
Db 876 TCCTTAACCTCAGACCTTACGACCATTAAGACTTCGCGAGAACTGAGAAGCC 931

RESULT 5
AP000682/c
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone CMB9-32A1 map 11q24, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
ACCESSION
AP000682
VERSION
AP000682.3 GI:9844967
KEYWORDS
HTG; HTGS; PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 127347)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 127,347 genomic DNA of 11q24
Published Only in Database (1999)
2 (bases 1 to 127347)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
AUTHORS
Submitted (08-NOV-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgc.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Aug 18, 2000 this sequence version replaced gi:8118870.
COMMENT

```

----- Genome Center  
 Center: RIKEN Genomic Sciences Center (GSC)  
 Center code: RIKEN  
 Web site: http://hgp.gsc.riken.go.jp/  
 Contact: hattori@gs.riken.go.jp  
 ----- Project Information  
 Center project name: HumDraft11  
 Center clone name: CMB9-32A1  
 ----- Summary Statistics

Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads  
 Assembly program: Phrap; version 0.930329  
 Consensus quality: 124981 bases at least Q40  
 Consensus quality: 125815 bases at least Q30  
 Consensus quality: 126286 bases at least Q20  
 Insert size: 126647; sum-of-contigs  
 Quality coverage: 12.71x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 33532 contig of 33532 bp in length  
 33633 59223 contig of 25591 bp in length  
 59324 59223 contig of 23684 bp in length  
 83108 100196 contig of 17089 bp in length  
 100297 112313 contig of 12017 bp in length  
 121414 121043 contig of 8630 bp in length  
 121144 125917 contig of 4774 bp in length  
 126018 127347 contig of 1330 bp in length.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 33532: contig of 33532 bp in length  
 \* 33633: gap of 100 bp  
 \* 59223: contig of 25591 bp in length  
 \* 59324 59223: gap of 100 bp  
 \* 83108 100196: contig of 23684 bp in length  
 \* 83108 83107: gap of 100 bp  
 \* 100197 100196: contig of 17089 bp in length  
 \* 100197 100296: gap of 100 bp  
 \* 100297 112313: contig of 12017 bp in length  
 \* 112314 121413: gap of 100 bp  
 \* 121414 121043: contig of 8630 bp in length  
 \* 121044 121143: gap of 100 bp  
 \* 121144 125917: contig of 4774 bp in length  
 \* 125918 126017: gap of 100 bp  
 \* 126018 127347: contig of 1330 bp in length.

#### FEATURES

Location/Qualifiers  
 1..127347  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11q24"  
 /clone="CMB9-32A1"

1..33532  
 /note="assembly\_fragment"  
 33633..59223  
 /note="assembly\_fragment"  
 59324..83007  
 /note="assembly\_fragment"  
 83108..100196  
 /note="assembly\_fragment clone\_end:SP6 vector\_side:left"  
 100297..112313

misc\_feature 112414..121043 /note="assembly\_fragment"  
 misc\_feature 121144..125917 /note="assembly\_fragment"  
 misc\_feature 126018..127347 /note="assembly\_fragment"  
 BASE COUNT 36325 a 27508 c 27825 g 34989 t 700 others  
 ORIGIN  
 Query Match 90.8%; Score 216; DB 2; Length 127347;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-53;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 21 AGGTGGCTCTGAACACCTGAGGAGCTGGACATCCCATGTTGACGATGTCAATGGCATCAG 80  
 Db |||||  
 Qy 81 GAGGGGCCCCAAGGCCCCCCTCCCTTCATGATGATTCATTTCTGTTTCATTCATT 140  
 Db |||||  
 Qy 8691 GAGGGGCCCCAAGGCCCCCCTCCCTTCATGATGATTCATTTCTGTTTCATTCATT 8632  
 Db |||||  
 Qy 141 CATCCATACATCCATCGCTCTGAGCTTTCACCTCTGACTCCCTCACTCCATCCATCAGACCT 200  
 Db |||||  
 Qy 8631 CATCCATACATCCATCGCTCTGAGCTTTCACCTCTGACTCCCTCACTCCATCAGACCT 8572  
 Db |||||  
 Qy 201 CTACGCCATCAAGACTCTGCCAGAACTGAGAAGCC 236  
 Db |||||  
 8571 CTACGCCATCAAGACTCTGCCAGAACTGAGAAGCC 8536  
 RESULT 6  
 AC063921  
 LOCUS AC063921 144833 bp DNA linear HTG 08-JAN-2003  
 DEFINITION Homo sapiens chromosome 3 clone RP11-142P10, WORKING DRAFT  
 SEQUENCE, 13 unordered pieces.  
 AC063921  
 ACCESSION AC063921.18 GI:2033587  
 VERSION HTG; HTGS PHASE1; HTGS\_DRAFT.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 144833)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
 Alsebrook, S.L., Amaral, H.C., Are, J.R., Ayele, M., Banks, R.,  
 Barbara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flegg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisedge, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mathew, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rojuben, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Tsanai, K., Vaquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wlezyk, R., Woodson, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 144833)  
Worley, K.C.

Direct Submission  
Submitted (22-APR-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 144833)  
Worley, K.C.

Direct Submission  
Submitted (08-JAN-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Apr 28, 2002 this sequence version replaced gi:16117924.

## COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: HAXY  
Center clone name: RP11-142P10  
----- Summary Statistics  
Sequencing vector: M13;  
Chemistry: Dye-terminator Big Dye; 14% of reads  
Assembly: Dye-terminator Big Dye; 86% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 127260 bases at least Q40  
Consensus quality: 132735 bases at least Q30  
Consensus quality: 136183 bases at least Q20  
Estimated insert size: 139078; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 13 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 2437: contig of 2437 bp in length  
2438 2537: gap of unknown length  
2538 4812: contig of 2275 bp in length  
4813 4912: gap of unknown length  
4913 8488: contig of 3576 bp in length  
8489 8598: gap of unknown length  
8599 11163: contig of 2575 bp in length  
11164 11263: gap of unknown length  
11264 16638: contig of 5375 bp in length  
16639 16739: gap of unknown length  
16739 21870: contig of 5132 bp in length  
21871 21970: gap of unknown length  
21971 30780: contig of 8810 bp in length  
30781 30880: gap of unknown length  
30881 39287: contig of 8387 bp in length  
39288 46890: gap of unknown length  
46891 46990: gap of unknown length

46991 59656: contig of 12666 bp in length  
59657 59756: gap of unknown length  
59757 83309: contig of 23553 bp in length  
83310 83409: gap of unknown length  
83410 111857: contig of 28448 bp in length  
111858 111857: gap of unknown length  
111859 144833: contig of 32876 bp in length

FEATURES  
Location/Qualifiers  
source  
1..144833  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-142P10"

BASE COUNT 38456 a 33920 c 34115 g 37093 t 1249 others  
ORIGIN

Query Match 90.8%; Score 216; DB 2; Length 144833;  
Best Local Similarity 100.0%; Pred. No. 6e-53; 0; Indels 0; Gaps 0;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 AGGTGGCTGAAACACCTGAGGACTGGACATCCCATGTTTCAGCAATGTAATGGCATCAG 80  
DB 90523 AGGTGGCTGAAACACCTGAGGACTGGACATCCCATGTTTCAGCAATGTAATGGCATCAG 90582

QY 81 GAGGGGCGCCCGAGGCGCCCATGCTCCCTTCATGCTCCATGCTCTGTTTCATTCATT 140  
DB 90583 GAGGGGCGCCCGAGGCGCCCATGCTCCCTTCATGCTCCATGCTCTGTTTCATTCATT 90642

QY 141 CATCCATACATCACCCTGCTGAGCTTTTCACCTCCCTGAGTCCCTCACTCCATCAGACT 200  
DB 90643 CATCCATACATCACCCTGCTGAGCTTTTCACCTCCCTGAGTCCCTCACTCCATCAGACT 90702

QY 201 CTAGGACCATTAAGACTCTGCCAGACTGAGAGCC 236  
DB 90703 CTAGGACCATTAAGACTCTGCCAGACTGAGAGCC 90738

RESULT 7  
AC021981 149800 bp DNA linear HTG 19-JUL-2000  
LOCUS Homo sapiens chromosome 11 clone RP11-11C15 map 11, WORKING DRAFT  
DEFINITION SEQUENCES, 29 unordered pieces.  
AC021981  
VERSION AC021981.2 GI:9280747  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 11, clone RP11-11C15  
Unpublished  
REFERENCE  
2 (bases 1 to 149800)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Becker, R., Bedalov, F.,  
Boguslavsky, L., Bouckhelter, B., Brown, A., Burkett, G., Castle, A.,  
Choe, Y., Collange, M., Collins, S., Collins, S., Collins, S., Cooke, P.,  
DeArnell, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,  
Lander, T., Lech, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
Maddaloni, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPherson, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,  
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, P., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodor, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.  
Direct Submission

TITLE

## JOURNAL

## COMMENT

Submitted (23-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 19, 2000 this sequence version replaced g1:6731265. All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIPR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L1335

Center clone name: 11.C.15

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 13517 bases at least Q40

Consensus quality: 141872 bases at least Q30

Consensus quality: 144404 bases at least Q20

Insert size: 121000; agarose-fp

Quality coverage: 4.2 in Q20 bases; agarose-fp

Quality coverage: 3.5 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 29 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1441: contig of 1441 bp in length
1 1541: gap of 100 bp
1 1542: 2883: contig of 1342 bp in length
1 2884: 2983: gap of 100 bp
1 2984: 4649: contig of 1666 bp in length
1 4650: 4749: gap of 100 bp
1 4750: 6112: contig of 1363 bp in length
1 6113: 6212: gap of 100 bp
1 6213: 7413: contig of 1201 bp in length
1 7414: 7513: gap of 100 bp
1 7514: 9050: contig of 1537 bp in length
1 9051: 9150: gap of 100 bp
1 9151: 10944: contig of 1794 bp in length
1 10945: 11044: gap of 100 bp
1 11045: 12203: contig of 2159 bp in length
1 12204: 13303: gap of 100 bp
1 13304: 16118: contig of 2815 bp in length
1 16119: 16218: gap of 100 bp
1 16219: 18673: contig of 2455 bp in length
1 18674: 18773: gap of 100 bp
1 18774: 21180: contig of 2407 bp in length
1 21181: 21280: gap of 100 bp
1 21281: 23200: contig of 1920 bp in length
1 23201: 23300: gap of 100 bp
1 23301: 26226: contig of 3226 bp in length
1 26227: 26626: gap of 100 bp
1 26627: 29655: contig of 3029 bp in length
1 29656: 29755: gap of 100 bp
1 29756: 34738: contig of 4983 bp in length
1 34739: 34838: gap of 100 bp
1 34839: 39394: contig of 4556 bp in length
1 39395: 39494: gap of 100 bp
1 39495: 43669: contig of 4175 bp in length
1 43670: 43769: gap of 100 bp
1 43770: 49948: contig of 6179 bp in length
1 49949: 50048: gap of 100 bp
1 50049: 55578: contig of 5530 bp in length
1 55579: 60559: gap of 100 bp
1 60560: 60559: contig of 5281 bp in length

```

```

* 60960 61059: gap of 100 bp
* 61060 66394: contig of 5335 bp in length
* 66395 66494: gap of 100 bp
* 66495 72947: contig of 6453 bp in length
* 72948 73047: gap of 100 bp
* 73048 80874: contig of 7827 bp in length
* 80875 80974: gap of 100 bp
* 80975 87847: contig of 6873 bp in length
* 87848 87947: gap of 100 bp
* 87949 92279: contig of 9332 bp in length
* 92280 97379: gap of 100 bp
* 97380 106037: contig of 8658 bp in length
* 106038 106137: gap of 100 bp
* 106138 114206: contig of 8069 bp in length
* 114207 114306: gap of 100 bp
* 114307 129830: contig of 15524 bp in length
* 129831 129930: gap of 100 bp
* 129931 149800: contig of 19870 bp in length.
FEATURES
      source
      1. 149800
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="11"
      /map="11"
      /clone="RP11-11C15"
      /clone_lib="RPC1-11 Human Male BAC"
      1. 1441
      /note="assembly_fragment"
      1542. 2883
      /note="assembly_fragment"
      2984. 4649
      /note="assembly_fragment"
      4750. 6112
      /note="assembly_fragment"
      6213. 7413
      /note="assembly_fragment"
      7514. 9050
      /note="assembly_fragment"
      9151. 10944
      /note="assembly_fragment"
      11045. 13203
      /note="assembly_fragment"
      13304. 16118
      /note="assembly_fragment"
      16219. 18673
      /note="assembly_fragment"
      clone_end:T7
      vector_side:right"
      18774. 21180
      /note="assembly_fragment"
      21281. 23200
      /note="assembly_fragment"
      23301. 26526
      /note="assembly_fragment"
      26627. 29655
      /note="assembly_fragment"
      29756. 34738
      /note="assembly_fragment"
      34839. 39394
      /note="assembly_fragment"
      39495. 43669
      /note="assembly_fragment"
      43770. 49948
      /note="assembly_fragment"
      50049. 55578
      /note="assembly_fragment"
      clone_end:SP6
      vector_side:right"
      55679. 60559
      /note="assembly_fragment"
      61060. 66394
      /note="assembly_fragment"

```

```
misc_feature 66495..72947
/note="assembly_fragment"
misc_feature 73048..80874
/note="assembly_fragment"
misc_feature 80975..87847
/note="assembly_fragment"
misc_feature 87948..97279
/note="assembly_fragment"
misc_feature 97380..106037
/note="assembly_fragment"
misc_feature 106138..114206
/note="assembly_fragment"
misc_feature 114307..129830
/note="assembly_fragment"
misc_feature 129931..149800
/note="assembly_fragment"
BASE COUNT 39569 a 33827 c 32852 g 40744 t 2808 others
ORIGIN
Query Match 90.8%; Score 216; DB 2; Length 149800;
Best Local Similarity 100.0%; Pred. No. 6e-53;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 AGGTGGCTGAAACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCATGGCATCAG 80
DB 37101 AGGTGGCTGAAACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCATGGCATCAG 37160
QY 81 GAGGGCGCCCAAGGGCCCATGCTCCCTTCATGTCATCCATTTCTTGTTCATTCATT 140
DB 37161 GAGGGCGCCCAAGGGCCCATGCTCCCTTCATGTCATCCATTTCTTGTTCATTCATT 37220
QY 141 CATCATACATCCACCTGCTCTGAGCTTTCACCTGTCATCCCTCACTCCATCCATCAGACCT 200
DB 37221 CATCATACATCCACCTGCTCTGAGCTTTCACCTGTCATCCCTCACTCCATCCATCAGACCT 37280
QY 201 CTAGCACCATAAGACTCTGCCAGACTGAGAGCC 236
DB 37281 CTAGCACCATAAGACTCTGCCAGACTGAGAGCC 37316
RESULT 8
AC024604/c 172546 bp DNA linear HTG 30-AUG-2001
LOCUS Homo sapiens chromosome 10 clone RP11-331G19, WORKING
DEFINITION DRAFT SEQUENCE, 22 unordered pieces.
ACCESSION AC024604
VERSION AC024604.3 GI:8389428
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 172546)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 172546)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Jun 9, 2000 this sequence version replaced gi:7549605.
----- Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
----- Project Information
Center project name: hg202
----- Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
```

Assembly program: Phrap; version 990315  
Consensus quality: 152033 bases at least Q40  
Consensus quality: 165535 bases at least Q30  
Consensus quality: 165256 bases at least Q20  
Insert size: 170446; sum-of-contigs  
Quality coverage: 4.3x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1339: contig of 1339 bp in length  
1439: gap of unknown length  
2789: contig of 1350 bp in length  
2889: gap of unknown length  
4562: contig of 1673 bp in length  
4663: gap of unknown length  
6344: contig of 1682 bp in length  
6444: gap of unknown length  
9221: contig of 2777 bp in length  
9321: gap of unknown length  
11997: contig of 2676 bp in length  
13097: gap of unknown length  
14867: contig of 2770 bp in length  
14868: gap of unknown length  
19919: contig of 4952 bp in length  
20019: gap of unknown length  
24125: contig of 4106 bp in length  
24225: gap of unknown length  
24226: gap of unknown length  
29444: contig of 5219 bp in length  
29445: gap of unknown length  
34568: contig of 5024 bp in length  
34569: gap of unknown length  
39414: contig of 4746 bp in length  
39415: gap of unknown length  
45885: contig of 6371 bp in length  
45886: gap of unknown length  
50986: contig of 5012 bp in length  
50987: gap of unknown length  
51097: gap of unknown length  
5143: contig of 8046 bp in length  
5243: gap of unknown length  
67779: contig of 8336 bp in length  
67879: gap of unknown length  
76429: contig of 8550 bp in length  
76529: gap of unknown length  
86704: contig of 10175 bp in length  
86804: gap of unknown length  
103147: contig of 16343 bp in length  
103148: gap of unknown length  
103247: gap of unknown length  
11343: contig of 13096 bp in length  
116443: gap of unknown length  
135318: contig of 18875 bp in length  
135418: gap of unknown length  
135319: gap of unknown length  
172546: contig of 37128 bp in length.

#### FEATURES

source

1. 172546  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="Chromosome 10"  
/clone\_lib="RP11-331G19"  
/clone\_lib="RP11-331G19"

BASE COUNT 47530 a 37712 c 38612 g 46556 t 2136 others  
ORIGIN

Query Match 90.8%; Score 216; DB 2; Length 172546;  
Best Local Similarity 100.0%; Pred. No. 6e-53;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 21 AGGTGGCTGAAACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCAATGGCATCAG 80
Db 170778 AGGTGGCTGAAACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCAATGGCATCAG 170719
QY 81 GAGGGGCCCCCAAGGCCCCCATCGCTTCCCTTCATGCATCCATTTCTGTTCATTCATT 140
Db 170718 GAGGGGCCCCCAAGGCCCCCATCGCTTCCCTTCATGCATCCATTTCTGTTCATTCATT 170659
QY 141 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 200
Db 170658 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 170599
QY 201 CTACGCACCATAGACTCTGCCAGAACTGAGAGCC 236
Db 170598 CTACGCACCATAGACTCTGCCAGAACTGAGAGCC 170563

RESULT 9
AP002765/c 178169 bp DNA linear PRI 15-MAR-2003
LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-634B22,
DEFINITION complete sequence.
ACCESSION AP002765
VERSION AP002765.3 GI:16751488
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Published Only in Database (2000)
2 (bases 1 to 178169)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Nov 5, 2001 this sequence version replaced gi:12381934.
FEATURES
Source 1..178169
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-634B22"

BASE COUNT 45062 a 39655 c 41140 g 51312 t
ORIGIN

Query Match 90.8%; Score 216; DB 9; Length 178169;
Best Local Similarity 100.0%; Pred. NO. 6e-53;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 AGGTGGCTGAAACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCAATGGCATCAG 80
Db 87137 AGGTGGCTGAAACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCAATGGCATCAG 87078
QY 81 GAGGGGCCCCCAAGGCCCCCATCGCTTCCCTTCATGCATCCATTTCTGTTCATTCATT 140
Db 87077 GAGGGGCCCCCAAGGCCCCCATCGCTTCCCTTCATGCATCCATTTCTGTTCATTCATT 87018
QY 141 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 200
Db 87017 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 86958
QY 201 CTACGCACCATAGACTCTGCCAGAACTGAGAGCC 236
Db 170598 CTACGCACCATAGACTCTGCCAGAACTGAGAGCC 170563

RESULT 10
AP002749/c 178431 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens chromosome 11 clone RP11-158J21 map 11q24, WORKING
DEFINITION DRAFT SEQUENCE, 41 unordered pieces.
ACCESSION AP002749
VERSION AP002749.1 GI:9188588
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Published Only in Database (2000)
2 (bases 1 to 178431)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
----- Genomic Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDrafc11
Center clone name: RP11-158J21
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 155982 bases at least Q40
Consensus quality: 165758 bases at least Q30
Consensus quality: 170239 bases at least Q20
Insert size: 174431; sum-of-contigs
Quality coverage: 4.20x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
41 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1 13012 contig of 13012 bp in length
13113 24677 contig of 11565 bp in length
24778 34640 contig of 9883 bp in length
34741 43136 contig of 8456 bp in length
43297 50126 contig of 6830 bp in length
50227 57725 contig of 7499 bp in length
57826 64612 contig of 6787 bp in length
64713 70999 contig of 6287 bp in length
71100 76188 contig of 5089 bp in length
76289 83580 contig of 7292 bp in length
83681 90511 contig of 6831 bp in length
90612 95493 contig of 4882 bp in length
95594 102262 contig of 6669 bp in length
102363 105866 contig of 3504 bp in length
105967 111142 contig of 5176 bp in length
111243 115923 contig of 4681 bp in length
116034 121747 contig of 5724 bp in length
121848 125532 contig of 3685 bp in length
125633 127927 contig of 2295 bp in length
```

128028 131184 contig of 3157 bp in length  
 131285 132725 contig of 1441 bp in length  
 132826 135566 contig of 2741 bp in length  
 135667 138000 contig of 2334 bp in length  
 138101 141194 contig of 3094 bp in length  
 141295 143619 contig of 2325 bp in length  
 143720 146448 contig of 2729 bp in length  
 146549 149531 contig of 2983 bp in length  
 149632 152157 contig of 2526 bp in length  
 152258 154933 contig of 2611 bp in length  
 155034 158334 contig of 2075 bp in length  
 158435 162343 contig of 1734 bp in length  
 162444 164809 contig of 2366 bp in length  
 164910 167136 contig of 2127 bp in length  
 167237 169169 contig of 2033 bp in length  
 169270 171052 contig of 1781 bp in length  
 171153 172933 contig of 1147 bp in length  
 173034 174180 contig of 1449 bp in length  
 174281 175729 contig of 1076 bp in length  
 175820 176905 contig of 1426 bp in length  
 176906 178431 contig of 1426 bp in length.  
 177006

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 41 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 13012: contig of 13012 bp in length  
 13013 13112: gap of 100 bp  
 13113 24677: contig of 11565 bp in length  
 24678 24777: gap of 100 bp  
 24778 34640: contig of 9863 bp in length  
 34641 34740: gap of 100 bp  
 34741 43196: contig of 8456 bp in length  
 43197 43296: gap of 100 bp  
 43297 50127: contig of 6830 bp in length  
 50128 50226: gap of 100 bp  
 50227 57726: contig of 7499 bp in length  
 57727 57825: gap of 100 bp  
 57826 64612: contig of 6787 bp in length  
 64613 64712: gap of 100 bp  
 64713 71099: contig of 6287 bp in length  
 71100 71099: gap of 100 bp  
 71100 76188: contig of 5089 bp in length  
 76189 83580: gap of 100 bp  
 83581 83680: contig of 7292 bp in length  
 83681 90511: gap of 100 bp  
 90512 90611: contig of 6831 bp in length  
 90612 95493: contig of 4882 bp in length  
 95494 95593: gap of 100 bp  
 102262 102262: contig of 6669 bp in length  
 102263 102362: gap of 100 bp  
 102363 105868: contig of 3504 bp in length  
 105869 105967: gap of 100 bp  
 105968 111142: contig of 5176 bp in length  
 111143 111243: gap of 100 bp  
 111244 115923: contig of 4681 bp in length  
 115924 116024: gap of 100 bp  
 116025 121747: contig of 5724 bp in length  
 121748 121847: gap of 100 bp  
 121848 125332: contig of 3685 bp in length  
 125333 125633: gap of 100 bp  
 125634 127927: contig of 2295 bp in length  
 127928 128027: gap of 100 bp  
 128028 131184: contig of 3157 bp in length  
 131185 131285: gap of 100 bp  
 131286 132726: contig of 1441 bp in length  
 132727 135566: contig of 2741 bp in length

135567 135666: gap of 100 bp  
 135667 138000: contig of 2334 bp in length  
 138001 138101: gap of 100 bp  
 138102 141194: contig of 3094 bp in length  
 141195 141295: gap of 100 bp  
 141296 143619: contig of 2325 bp in length  
 143720 146448: gap of 100 bp  
 146449 149531: contig of 2983 bp in length  
 149532 152157: contig of 2526 bp in length  
 152158 154933: contig of 2611 bp in length  
 154934 158334: contig of 2075 bp in length  
 158335 162343: contig of 1734 bp in length  
 162344 164809: contig of 2366 bp in length  
 164810 167136: contig of 2127 bp in length  
 167137 169169: contig of 2033 bp in length  
 169170 171052: contig of 1781 bp in length  
 171053 172933: contig of 1147 bp in length  
 172934 174180: contig of 1449 bp in length  
 174181 175729: contig of 1076 bp in length  
 175730 176905: contig of 1426 bp in length  
 176906 178431: contig of 1426 bp in length.  
 177006

FEATURES

source  
 1. 178431  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11q24"  
 /clone="RP11-158J21"  
 1. 13012  
 /note="assembly\_fragment"  
 13113..24677  
 /note="assembly\_fragment"  
 24778..34640  
 /note="assembly\_fragment"  
 34741..43196  
 /note="assembly\_fragment"  
 43297..50126  
 /note="assembly\_fragment"  
 50227..57725  
 /note="assembly\_fragment"  
 57826..64612

Query Match 90.8%; Score 216; DB 2; Length 178431;  
 Best Local Similarity 100.0%; Pred. No. 6e-53;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 AGGTGGCTGAACACCTGAGGAGCTGACATCCATGTCAGCAATGTCATGCTCAG 80  
 Db 173435 AGGTGGCTGAACACCTGAGGAGCTGACATCCATGTCAGCAATGTCATGCTCAG 173376  
 QY 81 GAGGGCGCCCAAGCGCCCGCCATGCTCCCTTCATGTCATGTCATGTCATGTCAT 140  
 Db 173375 GAGGGCGCCCAAGCGCCCGCCATGCTCCCTTCATGTCATGTCATGTCATGTCAT 173316  
 QY 141 CATCATACATCCACCTGCGCTCTGAGCTTTACCTCTGACTCCCTAACTCCATCAGACCT 200









```

51870 93650 contig of 41781 bp in length
93751 118302 contig of 24552 bp in length
118403 135796 contig of 73934 bp in length
135887 147076 contig of 11180 bp in length
147177 153868 contig of 6692 bp in length
153969 159683 contig of 5715 bp in length
159784 162363 contig of 2580 bp in length
162454 163651 contig of 1188 bp in length
163752 165836 contig of 1885 bp in length
165737 167342 contig of 1606 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	51769:	contig of 51769 bp in length
*	51869:	gap of 100 bp
*	51870:	contig of 41781 bp in length
*	93750:	gap of 100 bp
*	93751:	gap of 100 bp
*	118302:	contig of 24552 bp in length
*	118303:	118402: gap of 100 bp
*	118403:	135796: contig of 17394 bp in length
*	135797:	135896: gap of 100 bp
*	135798:	147076: contig of 11180 bp in length
*	147077:	147176: gap of 100 bp
*	147177:	153868: contig of 6692 bp in length
*	153869:	153968: gap of 100 bp
*	153969:	159683: contig of 5715 bp in length
*	159684:	159783: gap of 100 bp
*	159784:	162363: contig of 2580 bp in length
*	162364:	162463: gap of 100 bp
*	162464:	163551: contig of 1188 bp in length
*	163552:	163751: gap of 100 bp
*	163752:	165836: contig of 1885 bp in length
*	165837:	165736: gap of 100 bp
*	165737:	167342: contig of 1606 bp in length.

## FEATURES

[illegible]

### Query Match

18.5%; Score 44; DB 2; Length 167342;

	Best Local Similarity	63.0%;	Pred. No. 0.029;			
	Matches	68;	Conservative	0;	Mismatches	40; Indels 0; Gaps 0;
Qy	103	CGCTTCCTTCATGTCATCCCATGTTCTGTTCATTCAATCAATCCATACATCCACCTGCCTTC	162			
Db	162227	CGCATCCACCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAC	162168			
Qy	163	TGAGCTTTCACCTCTGACTCCCTTAACTCCCATCCAGACCTCTTACGACCA	210			
Db	162167	CCACCATCCACCCATTTATCATCCATTCCTCCCTCCCTCCATTCACCA	162120			

Search completed: December 5, 2003, 20:51:07  
Job time : 1000.27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 14:12:31 ; Search time 89.6511 Seconds  
(without alignments)  
7166.298 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_1024\_1261

Perfect score: 238

Sequence: 1 aacagagcagtgtgacatg.....tgccagaactgagaacggcgg 238

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1\_ Geneseq 19Jun03.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
26: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	238	100.0	1261	21	AAC67837
2	236	99.2	1510	22	AAF84146
3	236	99.2	4052	24	ABA93727
4	216	90.8	4625	22	ABL04971
5	216	90.8	4825	23	ABL97864
6	38.6	16.2	11089	22	AAK79653
7	38	16.0	38360	24	ABQ75124
8	37.8	15.9	3768	24	ABK64794

C 9	37.8	15.9	3768	24	ABL62894	Breast cancer rela
C 10	37.8	15.9	3768	24	ABL63103	Breast cancer rela
C 11	37.8	15.9	3768	24	ABL69598	Prostate cancer re
C 12	37.6	15.8	1367	21	AAAB673	N. meningitidis pa
C 13	37.4	15.7	1576	22	AAK73497	Human immune/haema
C 14	37.4	15.7	7954	22	AAK73498	Human immune/haema
C 15	37	15.5	2632	22	AAK90602	Rat sodium channel
C 16	37	15.5	3108	22	AAK90600	Rat sodium channel
C 17	37	15.5	6419	24	ABL32266	Human immune syste
C 18	36.6	15.4	4743	22	AAK05210	Human reproductive
C 19	36.6	15.4	4743	23	ABL98093	Human testicular a
C 20	36.6	15.4	4747	22	AAK05209	Human reproductive
C 21	36.6	15.4	4747	23	ABL98092	Human testicular a
C 22	35.8	15.0	168174	24	ABT11173	Human 5-lipoxygena
C 23	35.8	15.0	168273	24	ABT11114	Human immune syste
C 24	35.4	14.9	12142	24	ABL33672	Human immune syste
C 25	35.2	14.8	381	24	ABN24722	Human ORFX polynuc
C 26	35.2	14.8	758	22	AAI95421	Human neuroblastom
C 27	35.2	14.8	249487	24	ABN85733	Mouse genomic regi
C 28	34.6	14.5	6155	22	AAK40566	DNA encoding human
C 29	34.6	14.5	6155	22	AAK06454	Human reproductive
C 30	34.6	14.5	32249	22	ABA17155	Human nervous syst
C 31	34.6	14.5	185035	24	ABT10147	Human breast cance
C 32	34.4	14.5	5490	21	AAA38099	Human genomic DNA
C 33	34.4	14.5	6140	22	AAK05394	Human reproductive
C 34	34.4	14.5	6140	23	ABL98256	Human testicular a
C 35	34.4	14.5	32169	22	ABA14358	Human nervous syst
C 36	34.4	14.5	59747	24	ABQ88209	Human osteoblast d
C 37	34.4	14.5	148834	24	ABK83570	Human cDNA differe
C 38	34.2	14.4	32249	22	AAK04473	Human reproductive
C 39	34.2	14.4	49777	22	AAK75029	Human immune/haema
C 40	34.2	14.4	49777	22	AAK76214	Human immune/haema
C 41	34.2	14.4	567571	25	AAD53224	Human chromosome 3
C 42	34	14.3	488	22	ABA42824	Human breast cell
C 43	34	14.3	488	22	AAK01502	Human brain expres
C 44	34	14.3	488	22	AAK26947	Human bone marrow
C 45	34	14.3	488	22	AAI01483	Probe #1474 used t

## ALIGNMENTS

RESULT 1  
AAC67837  
ID AAC67837 standard; cDNA; 1261 BP.

AC AAC67837;

DT 15-FEB-2001 (first entry)

DE Human beta3 cDNA.

Human; beta sub-unit; beta3; analgesic; anticonvulsant;

cerebroprotective; vasotropic; cardiac; nootropic; cytostatic;

dermatological; gene therapy; voltage-gated sodium channel; pain;

epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;

familial nonchromaffin paraganglioma; phenylketonuria;

Charcot Marie Tooth disease; ss.

OS Homo sapiens.

FN WO2000063367-A1.

PD 26-OCT-2000.

PF 24-FEB-2000; 2000WO-EP01783.

PR 15-APR-1999; 99US-0129473.

PA (WARN ) WARNER LAMBERT CO.

PP (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

PI Cox P, Dixon A, Jackson A, Morgan K;

XX DR WPI: 2000-665241/64.  
XX DR P-PSDB; AAB36002.  
XX  
XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
PT sodium channel, and their corresponding polypeptides, useful for  
PT detecting and treating sodium channel-associated conditions, e.g. pain,  
PT epilepsy and stroke  
XX  
XX Claim 10; Page 70-71; 88pp; English.  
XX  
XX The present sequence is given in the claims of a specification  
CC relating to a novel family of beta sub-unit proteins from a  
CC voltage-gated sodium channel. Human and rat beta sub-units, which  
CC have been collectively identified as beta3, have been isolated.  
CC The polynucleotides and polypeptides are useful for screening for  
CC agonists and antagonists of sodium channels. The agonists, antagonists,  
CC proteins and nucleic acids may be used diagnosing of treating diseases  
CC or conditions associated with voltage-gated sodium channels, e.g. pain,  
CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial  
CC Nonchromaffin paraganglioma, Phenylketonuria and Charcot Marie Tooth  
CC disease.  
XX  
XX SQ Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 other;  
Query Match 100.0%; Score 238; DB 21; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 2e-69; Indels 0; Gaps 0;  
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGACCTGAAACCTGAGGAGCTGACATCCCATGTTTC 60  
Db 1024 AACAGGAGCAGTGTGACATGAGTGGCTGACCTGAAACCTGAGGAGCTGACATCCCATGTTTC 1083  
QY 61 AGCAATGTCAATGGCATCAGGAGGGGCGCCCAAGGGCCCGCATCGCTTCCCTTCATGTCATC 120  
Db 1084 AGCAATGTCAATGGCATCAGGAGGGGCGCCCAAGGGCCCGCATCGCTTCCCTTCATGTCATC 1143  
QY 121 CATGTGTTCTGTTCAATTCATTCATTCATCCATACATCCATCCCTGCTGAGCTTTCACCTCTGAC 180  
Db 1144 CATGTGTTCTGTTCAATTCATTCATTCATCCATACATCCATCCCTGCTGAGCTTTCACCTCTGAC 1203  
QY 181 TCCCTAACTCCATCAGACCTCTAGCAGCAGTACCTGCTGAGCTGAGAGCGG 238  
Db 1204 TCCCTAACTCCATCAGACCTCTAGCAGCAGTACCTGCTGAGCTGAGAGCGG 1261

RESULT 2  
AAF84146  
ID AAF84146 standard; cDNA; 1510 BP.  
XX  
XX AAF84146;  
AC  
XX  
XX 07-SEP-2001 (first entry)  
DT  
XX  
XX Human novel sodium channel beta1-like subunit encoding cDNA.  
DE  
XX  
XX Sodium channel; sensory neurone specific channel; beta1-like subunit;  
KW SNS; therapeutic; pain; analgesic; ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 213..860  
FT /tag= a  
FT /product= "sodium channel beta1-like subunit"  
FT  
XX WO200144293-A2.  
XX  
XX 21-JUN-2001.  
XX  
XX 14-DEC-2000; 2000WO-GB04802.  
XX  
XX 17-DEC-1999; 99GB-0029970.  
XX  
XX

XX (GLAX ) GLAXO GROUP LTD.  
XX  
XX Plumptre M, Powell AJ, Sanseau P;  
XX WPI: 2001-398129/42.  
XX P-PSDB; AAB85206.  
XX  
XX Novel sub-unit for voltage-gated sodium channel proteins for producing  
PT agents useful for treating pain  
XX  
XX Claim 4; Page 29-30; 31pp; English.  
XX  
XX The invention provides a novel beta1-like sub-unit for voltage-gated  
CC sodium ion channel polypeptide, specifically a sensory neurone specific  
CC channel (SNS) subunit. The novel beta1-like subunit is useful for  
CC producing a therapeutic agent which is useful treating pain in a patient.  
CC The subunit can be expressed by standard recombinant methodology. The  
CC present sequence represents a human novel sodium channel beta1-like  
CC subunit encoding cDNA.  
XX  
XX SQ Sequence 1510 BP; 350 A; 410 C; 414 G; 336 T; 0 other;  
Query Match 99.2%; Score 236; DB 22; Length 1510;  
Best Local Similarity 100.0%; Pred. No. 9.8e-69;  
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGACCTGAAACCTGAGGAGCTGACATCCCATGTTTC 60  
Db 861 AACAGGAGCAGTGTGACATGAGTGGCTGAAACCTGAGGAGCTGACATCCCATGTTTC 920  
QY 61 AGCAATGTCAATGGCATCAGGAGGGGCGCCCAAGGGCCCGCATCGCTTCCCTTCATGTCATC 120  
Db 921 AGCAATGTCAATGGCATCAGGAGGGGCGCCCAAGGGCCCGCATCGCTTCCCTTCATGTCATC 980  
QY 121 CATGTGTTCTGTTCAATTCATTCATTCATCCATACATCCCTGCTGAGCTTTCACCTCTGAC 180  
Db 981 CATGTGTTCTGTTCAATTCATTCATTCATCCATACATCCCTGCTGAGCTTTCACCTCTGAC 1040  
QY 181 TCCCTAACTCCATCAGACCTCTAGCAGCAGTACCTGCTGAGCTGAGAGCGG 236  
Db 1041 TCCCTAACTCCATCAGACCTCTAGCAGCAGTACCTGCTGAGCTGAGAGCGG 1096

RESULT 3  
ABA93727  
ID ABA93727 standard; cDNA; 4052 BP.  
XX  
XX ABA93727;  
AC  
XX  
XX 30-APR-2002 (first entry)  
DT  
XX  
XX Human signal transduction cDNA clone amy2\_2f18.  
DE  
XX  
XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;  
KW gene therapy; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200198454-A2.  
XX  
XX 27-DEC-2001.  
XX  
XX 25-APR-2001; 2001WO-IB02050.  
XX  
XX 25-APR-2000; 2000US-199380P.  
XX  
XX (GSHU-) GERMAN HUMAN GENOME PROJECT.  
XX  
XX Wiemann S;  
XX  
XX WPI: 2002-055860/07.  
XX P-PSDB; ABB05689.  
XX

```
XX Human cDNA sequences and clones derived from human fetal brain, fetal
PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
PT screening and therapy -
XX Claim 1; Page 174-175; 61pp; English.
XX The present invention describes assemblages and computer readable media
CC comprising novel human cDNA sequences and clones derived from human
CC fetal brain, foetal kidney, melanoma, testis and amygdala cDNA
CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
CC present invention which encode the proteins given in ABB05662 to
CC ABB05729. The human cDNA sequences and clones can be used in gene
CC therapy. The clones may be used in a variety of applications, for
CC example they may be used in profiling assays, for providing large arrays
CC of human genetic material for implementing large-scale screening
CC strategies and for treating diseases via gene therapy procedures.
XX
SQ Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 other;
Query Match 99.2%; Score 236; DB 24; Length 4052;
Best Local Similarity 100.0%; Pred. No. 1.5e-68;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGAGCAGTGTGACATGAGTGGCTGGAACACTGAGGACTGGACATCCCATGTC 60
Db 1452 AACAGAGCAGTGTGACATGAGTGGCTGGAACACTGAGGACTGGACATCCCATGTC 1511
QY 61 ACCATGTCATGAGCAGTGGAGGGCGCCCGAGGCGCCCATCGCTTCCCTTCATGTCATC 120
Db 1512 AGCAATGTCATGAGCAGTGGAGGGCGCCCGAGGCGCCCATCGCTTCCCTTCATGTCATC 1571
QY 121 CATTTGTTCTGTTCAATTCATTCATCATCATCATCCACTGCTCTGAGCTTTACCTCTGAC 180
Db 1572 CATTTGTTCTGTTCAATTCATTCATCATCATCATCCACTGCTCTGAGCTTTACCTCTGAC 1631
QY 181 TCCCTTAACCTCAGACTGAGCTTACGACCATGAGACTTGCAGAACTGAGAGGCC 236
Db 1632 TCCCTTAACCTCAGACTGAGCTTACGACCATGAGACTTGCAGAACTGAGAGGCC 1687
RESULT 4
AAL04971
ID AAL04971 standard; DNA; 4625 BP.
AC AAL04971;
XX
XX 21-NOV-2001 (first entry)
DE Human reproductive system related antigen DNA SEQ ID NO: 7659.
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX Homo sapiens.
XX WO200155320-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
```



21 AGGTGGCGTGAACACCTGAGGAGCTGGACATCCCATGTTTCAGCAATGTCAATGGCATCAG 80  
Db  
33 AGGTGGCGTGAACACCTGAGGAGCTGGACATCCCATGTTTCAGCAATGTCAATGGCATCAG 92  
QY 81 GAGGGCGCCCAAGGGCCCCATCGCTTCCCTTCATGCATCCATTGTTCTGTTCAATTCATT 140  
Db 93 GAGGGCGCCCAAGGGCCCCATCGCTTCCCTTCATGCATCCATTGTTCTGTTCAATTCATT 152  
QY 141 CATCCATACATCCCACTGCGCTCTGAGGTTTTACCTCTGACTCCCTAACTCCATCAGACCT 200  
Db 153 CATCCATACATCCCACTGCGCTCTGAGGTTTTACCTCTGACTCCCTAACTCCATCAGACCT 212  
QY 201 CTACGCGCCATAGACTCTCTGCCAGAACTGAGAGCC 236

Db 213 CTACGACCATAGACTCTCCAGACTGAGAGCC 248

RESULT 6  
AAK79653  
ID AAK79653 standard; DNA; 11089 BP.  
XX AC AAK79653;  
XX DT 07-NOV-2001 (first entry)  
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34465.  
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX KN Cytostatic; Gene therapy; vaccine; metastasis; ds.  
XX OS Homo sapiens.  
XX PN WC200157182-A2.  
XX PD 09-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01354.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0231415.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and  
XX metastasis.  
XX  
XX Disclosure; SEQ ID NO 3445; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
XX activity, and can be used in gene therapy and vaccine production. (I)  
XX proteins and polynucleotides may be used in the prevention, diagnosis and  
XX treatment of diseases associated with inappropriate (I) expression. For  
XX example, they may be used to treat disorders associated with decreased  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of (I) by expressing inactive proteins or to  
XX supplement the patients own production of (I). Additionally, (I)  
XX polynucleotides may be used to produce the secreted (I) by inserting  
XX the nucleic acids into a host cell and culturing the cell to express the  
XX protein. (I) proteins and polynucleotides may be used to prevent,  
XX diagnose and treat immune/haematopoietic-related diseases, especially  
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
XX to AAK87694 represent human immune/haematopoietic antigen genomic  
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
XX represent sequences used in the exemplification of the present invention.  
XX  
XX SQ Sequence 11089 BP; 2318 A; 3415 C; 3048 G; 2308 T; 0 other;  
Query Match 16.2%; Score 38.6; DB 22; Length 11089;  
Best Local Similarity 59.6%; Pred. No. 0.033;  
Matches 65; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
QY 105 CTTCCCTCATGCAATCCATGTTCTGTTTCATTCATTCATCCATCCATCCATCCATCC 164  
Db 1332 CTTCCATCATCCATCCATCCATTCATTCATCCATCCATCCATCCATCCATCCATCC 1391  
QY 165 AGCTTTCACCTGACCTCCCTAACTCCATGACCTCTAGCACCAATAA 213  
Db 1392 ATCCATCATCCATCCATTCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 1440  
RESULT 7  
ABQ75124  
ID ABQ75124 standard; DNA; 38360 BP.  
XX  
XX AC ABQ75124;  
XX  
XX DT  
XX 04-NOV-2002 (first entry)  
XX Human p15 region of chromosome 11 DNA sequence SEQ ID NO:14.  
XX Human; p15 region; chromosome 11; tumour growth; infertility; cytostatic;  
XX hepatitis C virus infection; antiinfertility; gene therapy; gene; ds.  
XX

OS Homo sapiens.  
XX  
XX WO200261085-A2.  
XX  
XX 08-AUG-2002.  
XX  
XX 31-OCT-2001; 2001WO-US45381.  
XX  
XX 31-OCT-2000; 2000US-244705P.  
XX  
XX (RYAN/) RYAN J W.  
XX Ryan JW;  
XX  
XX WPI; 2002-619251/66.  
XX  
XX New genes obtainable from the p15 region of human chromosome 11 (e.g.  
XX human achaete-scute homolog 2), useful in gene therapy, particularly  
XX for preventing or treating tumor growth, infertility or hepatitis C  
XX virus infection.  
XX  
XX Claim 21; Fig 3; 9app; English.  
XX  
XX The present invention describes an isolated genomic polynucleotide (I),  
XX which is obtainable from the p15 region of human chromosome 11.  
XX Specifically described are the human achaete-scute homologue 2 (HASH2),  
XX SMS3, tumour suppressing sub-transferable candidate 6 (TSSC6), ribosomal  
XX protein L26 (RIBO26), cluster of differentiation antigen 81 (CD 81) and  
XX tumour suppressing sub-transferable candidate 4 (TSSC4) genes as given  
XX in ABQ75117 to ABQ75122, encoding the proteins given in APPS2841 to  
XX APPS2846. Also described is an isolated polynucleotide obtainable from  
XX the p15 region of chromosome 11 having the sequence, which comprises any  
XX of two nucleotide sequences given in ABQ75123 and ABQ75124. (I) has  
XX cytostatic and antiinfertility activities and can be used in gene  
XX therapy. The HASH2, SMS3, TSSC6, RIBO26, CD 81 or TSSC4 polynucleotide  
XX sequences can be used for preventing, treating or ameliorating a medical  
XX condition, or for manufacturing a gene therapy for the prevention,  
XX treatment or amelioration of such medical conditions. An antisense  
XX oligonucleotide or mimetic of (I) can also be useful for manufacturing a  
XX medicament for the prevention, treatment or amelioration of these medical  
XX conditions. These conditions include tumour growth, infertility, or  
XX hepatitis C virus infection.  
XX  
XX SQ Sequence 38360 BP; 9681 A; 9897 C; 9285 G; 9497 T; 0 other;  
Query Match 16.0%; Score 38; DB 24; Length 38360;  
Best Local Similarity 62.8%; Pred. No. 0.087;  
Matches 59; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 97 CCCATCGCTTCCCTTCATGCAATCCATTCATTCATTCATTCATTCATTCATTCATTCAT 156  
Db 25309 CCCCTTCTTCTTCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 156  
QY 157 TGCTCTGAGCTTTCACCTCTGACTCCCTAACTC 190  
Db 25369 TATCTATCTAGTACCGAGCTAGCTACCTGTATC 25402  
RESULT 8  
ABK64794/c  
ID ABK64794 standard; DNA; 3768 BP.  
XX  
XX AC ABK64794;  
XX  
XX DT  
XX 18-JUN-2002 (first entry)  
XX Human benign prostatic hyperplasia gene #689.  
XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
XX  
XX OS  
XX Homo sapiens.  
XX  
XX WO200212440-A2.  
XX

XX PD 14-FEB-2002.

XX PF 07-AUG-2001; 2001WO-US24708.

XX PR 07-AUG-2000; 2000US-223323P.

XX PR 05-JUN-2001; 2001US-0873319.

XX PR (GENE-) GENE LOGIC INC.

XX PA (NISB ) JAPAN TOBACCO INC.

XX XX Munger WE, Kulxani P, Getzenberg RH, Waga I, Yamamoto J;

XX XX WPI; 2002-257476/30.

XX PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by

XX PT detecting expression levels of one or more genes in prostate cells from

XX PT patient that are differentially regulated compared to normal prostate

XX PT cells -

XX PS Disclosure; Page 371-372; 444pp; English.

XX CC The invention relates to a method of diagnosing (I) the onset or

XX CC progression of benign prostatic hyperplasia (BPH), or screening (II) for

XX CC or identifying an agent that modulates the onset or progression of BPH.

XX CC The method is based on changes in gene expression in BPH tissue isolated

XX CC from patients exhibiting different clinical states of prostate

XX CC hyperplasia as compared to normal prostate tissue. (I) comprises

XX CC detecting the expression levels of one or more genes in prostate cells

XX CC from the subject that are differentially regulated compared to normal

XX CC prostate cells. (II) comprises preparing a first gene expression profile

XX CC of BPH cells or BPH-like cell population, exposing the cells to the

XX CC agent, preparing a second gene expression profile of the agent exposed

XX CC cells, and comparing the first and second gene expression profiles.

XX CC (I) is useful for diagnosing the onset or progression of BPH. (II) is

XX CC useful for identifying an agent that modulates the onset or progression

XX CC of BPH. The methods are useful to present information identifying

XX CC the expression level in a tissue or cells, by comparing the expression

XX CC level of genes given in the specification in the tissue or cells to the

XX CC level of expression of gene in the database, and displaying the

XX CC expression levels of at least one gene in the tissue or cell sample

XX CC compared to the expression level in BPH. Agents using (II) are useful for

XX CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human

XX CC benign prostatic hyperplasia gene sequences of the invention.

XX SQ Sequence 3768 BP; 1030 A; 730 C; 1089 G; 919 T; 0 other;

Query Match 15.9%; Score 37.8; DB 24; Length 3768;

Best Local Similarity 60.0%; Pred. No. 0.039;

Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 105 CTTCCCTTCATGCATCCATGTTCTGTTTCATTCATTCATCCATCCATCCACCTGCTCTG 164

Db 1112 CATCCCTTCATTCATCCATCCACCTCCATCCCTTCATTCATTCATTCATTCCTCATTC 1053

QY 165 AGCTTCACCTCTGACTCCCTCACTCACTCCATCCATCCAGACCTCTACGCACC 209

Db 1052 ATCATCCCTCCCTCCATCCATCTATCCATCCATCCATCCACCCACCCATC 1008

RESULT 9

ABL62894/C

ID ABL62894 standard; DNA; 3768 BP.

XX AC ABL62894;

XX AC ABL62894;

XX DT 15-MAY-2002 (first entry)

XX DE Breast cancer related gene sequence SEQ ID NO:1231.

XX XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX KW Stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

KW gene; de.

XX OS Homo sapiens.

XX PN WO200194629-A2.

XX PD 13-DEC-2001.

XX PF 30-MAY-2001; 2001WO-US10838.

XX PR 05-JUN-2000; 2000US-209473P.

XX PR 05-JUN-2000; 2000US-209531P.

XX PR 18-SEP-2000; 2000US-233133P.

XX PR 18-SEP-2000; 2000US-233617P.

XX PR 20-SEP-2000; 2000US-234009P.

XX PR 20-SEP-2000; 2000US-234034P.

XX PR 20-SEP-2000; 2000US-234052P.

XX PR 22-SEP-2000; 2000US-234509P.

XX PR 22-SEP-2000; 2000US-234567P.

XX PR 25-SEP-2000; 2000US-234923P.

XX PR 25-SEP-2000; 2000US-234924P.

XX PR 25-SEP-2000; 2000US-235077P.

XX PR 25-SEP-2000; 2000US-235082P.

XX PR 25-SEP-2000; 2000US-235134P.

XX PR 25-SEP-2000; 2000US-235280P.

XX PR 26-SEP-2000; 2000US-235637P.

XX PR 26-SEP-2000; 2000US-235638P.

XX PR 27-SEP-2000; 2000US-235711P.

XX PR 27-SEP-2000; 2000US-235720P.

XX PR 27-SEP-2000; 2000US-235840P.

XX PR 28-SEP-2000; 2000US-235863P.

XX PR 28-SEP-2000; 2000US-236028P.

XX PR 28-SEP-2000; 2000US-236032P.

XX PR 28-SEP-2000; 2000US-236033P.

XX PR 28-SEP-2000; 2000US-236034P.

XX PR 28-SEP-2000; 2000US-236109P.

XX PR 28-SEP-2000; 2000US-236111P.

XX PR 29-SEP-2000; 2000US-236842P.

XX PR 29-SEP-2000; 2000US-236891P.

XX PR 02-OCT-2000; 2000US-237172P.

XX PR 02-OCT-2000; 2000US-237173P.

XX PR 02-OCT-2000; 2000US-237278P.

XX PR 02-OCT-2000; 2000US-237294P.

XX PR 02-OCT-2000; 2000US-237295P.

XX PR 02-OCT-2000; 2000US-237316P.

XX PR 03-OCT-2000; 2000US-237425P.

XX PR 03-OCT-2000; 2000US-237598P.

XX PR 03-OCT-2000; 2000US-237604P.

XX PR 03-OCT-2000; 2000US-237606P.

XX PR 03-OCT-2000; 2000US-237608P.

XX PR 01-NOV-2000; 2000US-244867P.

XX PR 01-NOV-2000; 2000US-245084P.

XX PA (AVAL-) AVALON PHARM.

XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

XX PI Soppet DR, Weaver Z;

XX DR WPI; 2002-188264/24.

XX PT Screening for anti-neoplastic agent involves exposing cells to a

XX PT chemical agent to be tested for anti-neoplastic activity, and

XX PT determining a change in expression of a gene of a signature gene set -

XX PS Claim 1; SEQ ID 1231; 44pp; English.

XX CC The present invention describes a method (M1) for screening for an

XX CC anti-neoplastic agent. The method involves exposing cells to a chemical

XX CC agent to be tested for anti-neoplastic activity, determining a change in

XX CC expression of at least one gene (I) of a signature gene set, where (I)

XX CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664

XX CC to ABL70110), or is at least 95% identical to (S), where a change in

XX CC expression is indicative of anti-neoplastic activity. (I) has cytostatic

activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, esophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 3768 BP; 1030 A; 730 C; 1089 G; 919 T; 0 other; XX

Sequence 3768 BP; 1030 A; 730 C; 1089 G; 919 T; 0 other;

Query Match	15.9%	Score 37.8	DB 24	Length 3768
Best Local Similarity	60.0%	Pred No. 0.039	42	Indels 0
Matches 63	Conservative	0	Mismatches	0
Gaps	0			

  

105	CTTCCCTTCATGCAATCATGTTCTGTTCAATTCATTCATCATCATCCACCTGCCTCTG	164
1112	CATCCCTCCCATTCATCATCCACCGCATCCCTTCATTCATTCATTCATTCCTCCATTC	1053
165	AGCTTTTCACTCTGACTCCCTAACTCCATCCATCAGACCTCTACGCAC	209
1052	ATCATCCCTCCCTCCATCCATTCATTCATCCATCACCCAGCCATC	1008

RESULT 10	
ABL63103/c	
ID	ABL63103 standard; DNA; 3768 BP.
XX	
XX	
AC	ABL63103;
XX	
XX	
DT	15-MAY-2002 (first entry)
XX	
XX	
DE	Breast cancer related gene sequence SEQ ID NO:1440.
KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW	cystostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW	gene; ds.
XX	
XX	
OS	Homo sapiens.

28-SEP-2000; 2000US-236034P  
PR  
28-SEP-2000; 2000US-236109P  
PR  
28-SEP-2000; 2000US-236111P  
PR  
29-SEP-2000; 2000US-236842P  
PR  
29-SEP-2000; 2000US-236851P  
PR  
02-OCT-2000; 2000US-237172P  
PR  
02-OCT-2000; 2000US-237173P  
PR  
02-OCT-2000; 2000US-237278P  
PR  
02-OCT-2000; 2000US-237294P  
PR  
02-OCT-2000; 2000US-237295P  
PR  
02-OCT-2000; 2000US-237316P  
PR  
03-OCT-2000; 2000US-237425P  
PR  
03-OCT-2000; 2000US-237598P  
PR  
03-OCT-2000; 2000US-237604P  
PR  
03-OCT-2000; 2000US-237608P  
PR  
03-OCT-2000; 2000US-237609P  
PR  
01-NOV-2000; 2000US-244867P  
PR  
01-NOV-2000; 2000US-245084P  
PR  
(AVAL-) AVALON PHARM.  
XX

Screening for anti-neoplastic agent involves exposing cells to a  
chemical agent to be tested for anti-neoplastic activity, and  
determining a change in expression of a gene of a signature gene set  
Claim 1: SEO ID 1440: 44pp: English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer. adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, bacillary carcinoma and Wilms' tumour.

Sequence 3768 BP: 1030 A; 730 C; 1089 G; 919 T; 0 other:

Query Match 15.9%; Score 37.8; DB 24; Length 3768;  
Best Local Similarity 60.0%; Pred. No. 0.039;  
Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0

RESULT 11	
ABL69598/c	
ID	ABL69598 standard; DNA; 3768 BP.
XX	
XX	ABL69598;
XX	
XX	
DT	15-MAY-2002 (first entry)
DT	
XX	



PR	30-JUN-2000	2000US-0215135
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216880
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	18-AUG-2000	2000US-0226759
PR	18-AUG-2000	2000US-0226779
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	23-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227009
PR	30-SEP-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	21-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234274
PR	25-SEP-2000	2000US-0234598
PR	25-SEP-2000	2000US-0234599
PR	25-SEP-2000	2000US-0234977
PR	26-SEP-2000	2000US-0235484
PR	26-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235836
PR	27-SEP-2000	2000US-0236327
PR	28-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	30-OCT-2000	2000US-0236370
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	13-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786

OS	Homo sapiens.
XX	
XX	
PN	WO200157182-A2.
XX	
XX	
PD	09-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01354.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180528.
PR	24-FEB-2000; 2000US-0184564.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.



PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226779.  
PR 22-AUG-2000; 2000US-0226861.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229387.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241825.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246509.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2001US-02559678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PI metastasis -  
XX  
PS Disclosure; SEQ ID NO 28310; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 7954 BP; 2012 A; 2128 C; 1983 G; 1831 T; 0 other;

Query Match 15.7%; Score 37.4; DB 22; Length 7954;  
Best Local Similarity 64.4%; Pred. No. 0.073;  
Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:08:07 ; Search time 703.619 Seconds  
(without alignments)  
8221.015 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_1024\_1261

Perfect score: 238

Sequence: 1 aacgggagcagtgtacatg.....tgccagaactgagaagccgg 238

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estlin.\*

4: em\_estma.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_hc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_eston.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_xod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gssi.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	236	99.2	977	13	EX452112
2	214.4	90.1	443	14	D44825
c	158.6	66.6	926	13	EX411288
4	101.2	42.5	484	14	CB441198

5	98.8	41.5	1201	13	EX420015
6	42.2	17.7	4149	11	AK049286
c	42	17.6	322	10	BG001495
c	41.8	17.6	601	14	CB442219
c	40.8	17.1	941	29	CNS02DGR
c	40.2	16.9	844	29	CNS02UEH
11	40	16.8	611	28	BH308136
12	39.2	16.5	676	28	AQ835006
c	39	16.4	564	28	AZ627327
c	38.6	16.2	660	14	BY730622
c	38.6	16.2	1101	29	CNS05122
c	38.4	16.1	668	29	AG887943
17	38.4	16.1	750	28	BH356347
18	38.2	16.1	751	28	AZ957360
19	38	16.0	835	28	AQ748173
20	38	16.0	1201	9	AL564828
21	37.8	15.9	963	13	EX328666
c	37.8	15.9	982	29	CNS05CTG
c	37.8	15.9	1045	29	CNS05M4G
24	37.6	15.8	436	28	AQ813992
25	37.6	15.8	445	28	AQ201281
26	37.4	15.7	372	10	BE181389
c	37.4	15.7	542	28	BZ173993
28	37.2	15.6	505	13	BZ258128
29	37.2	15.6	530	29	BZ689975
c	37.2	15.6	789	29	CNS02XIV
31	37.2	15.6	808	29	EX141482
c	37.2	15.6	852	29	CNS03HMZ
33	37.2	15.6	1258	10	BE035459
34	37	15.5	280	9	AA686211
c	37	15.5	304	29	EX166193
c	37	15.5	349	29	EX152709
c	37	15.5	378	29	EX240490
c	37	15.5	640	29	EX171544
c	37	15.5	666	29	EX232555
c	37	15.5	800	13	BU701454
41	37	15.5	802	29	EX187506
c	37	15.5	833	29	BZ284733
43	36.8	15.5	525	28	AQ336733
c	36.8	15.5	713	29	EX172652
c	36.8	15.5	799	29	EX200998

## ALIGNMENTS

RESULT 1

EX452112

LOCUS

DEFINITION

EX452112

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX452112 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
CS0DF023YA09 5-PRIME, mRNA sequence.  
EX452112 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
EX452112.1 GI:31024339  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 977)  
Li W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6147.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAG018ZF03 CS01654\_1&cluster=6147.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

```
Paradise Avenue Genoscope sequence ID : CSOBAG018ZF03_CS01654_1.
FEATURES
source
  Location/Qualifiers
    1..977
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DF023YA09"
      /tissue_type="FETAL BRAIN"
      /dev_stage="fetal"
      /clone_lib="Homo sapiens FETAL BRAIN"
      /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT      259 a 249 c 214 g 254 t 1 others
ORIGIN
Query Match      99.2%; Score 236; DB 13; Length 977;
Best Local Similarity 100.0%; Pred. No. 8.2e-56;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGGAGCAGTGACATGAGTGGCGCTGACACCTGAGGAGCTGACATCCCATGTTTC 60
Db 118 AACAGGAGCAGTGACATGAGTGGCGCTGACACCTGAGGAGCTGACATCCCATGTTTC 177
QY 61 AGCAATGTCATGAGGAGGCGCCCAAGGCGCCCATGCTTCCCTTCATGCATC 120
Db 178 AGCAATGTCATGAGGAGGCGCCCAAGGCGCCCATGCTTCCCTTCATGCATC 237
QY 121 CATTTGTTCTGTCATTCATTCATCCATGACATCCATGACCTGCTGAGCTTTCACCTCTGAC 180
Db 238 CATTTGTTCTGTCATTCATTCATCCATGACATCCATGACCTGCTGAGCTTTCACCTCTGAC 297
QY 181 TCCTTAATCCATGAGCAGCTTACGACACCATGAGCTTCCGAGACTGAGAGCC 236
Db 298 TCCTTAATCCATGAGCAGCTTACGACACCATGAGCTTCCGAGACTGAGAGCC 353
RESULT 2
D44825
LOCUS      HUNS0FY279 Human brain cDNA Homo sapiens cDNA clone NF220-K, mRNA
DEFINITION      sequence.
ACCESSION      D44825
VERSION      D44825.1 GI:1572300
KEYWORDS      EST.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  Hadano, S., Ishida, Y., Tomiyasu, H., Yamamoto, K., Bates, G.P. and
  Ikeda, J.
  Transcript map of the human chromosome 4p16.3 consisting of 627
  cDNA clones derived from 1 Mb of the Huntington's disease locus
  DNA Res. 3 (4), 239-255 (1996)
  97101646
  8946164
  Contact: Shinji Hadano
  Japan Science and Technology Corporation, NeuroGenes Project, ICORP
  Univ. of Tokai School of Med.
  Bohseidai, Isehara, Kanagawa 259-1193, Japan
  Tel: 81-463-91-5095
  Fax: 81-463-91-4993
  Email: shinji@ga.med.u-tokai.ac.jp.
  Location/Qualifiers
    1..443
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="NF220-K"
      /tissue_type="brain"
FEATURES
source
  Location/Qualifiers
    1..926
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DF023YA09"
      /tissue_type="FETAL BRAIN"
      /dev_stage="fetal"
      /clone_lib="Homo sapiens FETAL BRAIN"
      /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT      240 a 194 c 241 g 249 t 2 others
ORIGIN
Query Match      66.6%; Score 158.6; DB 13; Length 926;
Best Local Similarity 87.4%; Pred. No. 5.2e-34;
Matches 208; Conservative 0; Mismatches 25; Indels 5; Gaps 3;
```

```
/clone lib="Human brain cDNA"
BASE COUNT      111 a 130 c 82 g 120 t
ORIGIN
Query Match      90.1%; Score 214.4; DB 14; Length 443;
Best Local Similarity 99.5%; Pred. No. 7e-50;
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 AGTGGCCTGACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCATGGCATCAG 80
Db 33 AGTGGCCTGACACCTGAGGAGCTGACATCCATGTTTCAGCAATGTCATGGCATCAG 92
QY 81 GAGGCGCCCAAGGCGCCCATGCTTCCCTTCATGATCCATGTTCTGTTTCATTCATT 140
Db 93 GAGGCGCCCAAGGCGCCCATGCTTCCCTTCATGATCCATGTTCTGTTTCATTCATT 152
QY 141 CATCCATACATCCATGCTGCTGAGCTTTCACCTCTGACCTCCCTAACTCCATCAGACCT 200
Db 153 CATCCATACATCCATGCTGCTGAGCTTTCACCTCTGACCTCCCTAACTCCATCAGACCT 212
QY 201 CTACGACCATGAGACTCTGCCAGAACTGAGAAGCC 236
Db 213 CTACGACCATGAGACTCTGCCAGAACTGAGAAGCC 248
RESULT 3
BX411288/c
LOCUS      BX411288 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION      CS0DF023YA09 3-PRIME, mRNA sequence.
ACCESSION      BX411288
VERSION      BX411288.1 GI:30765121
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
  Full-length cDNA libraries and normalization
  Unpublished
  Contact: Genoscope
  Genoscope - Centre National de Sequencage
  BP 191 91006 EVRY cedex - France
  Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
  Library was constructed by Life Technologies, a division of
  Invitrogen. This sequence belongs to sequence cluster 6147.r For
  more information about this cluster, see
  http://www.genoscope.cns.fr/
  cgi-bin/cluster.cgi?seq=CS0BA10122D01_CS01060_1&cluster=6147.r.
  Contact : Feng Liang Email : fliang@lifetech.com URL :
  http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
  Faraday Avenue Genoscope sequence ID : CS0BA10122D01_CS01060_1.
  Location/Qualifiers
    1..926
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DF023YA09"
      /tissue_type="FETAL BRAIN"
      /dev_stage="fetal"
      /clone_lib="Homo sapiens FETAL BRAIN"
      /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT      240 a 194 c 241 g 249 t 2 others
ORIGIN
Query Match      66.6%; Score 158.6; DB 13; Length 926;
Best Local Similarity 87.4%; Pred. No. 5.2e-34;
Matches 208; Conservative 0; Mismatches 25; Indels 5; Gaps 3;
```

```

QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGCAACACCTGAGGAGCTGACATCCCATGTC 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
909 AACAGGAGCAGTGTG-CATAAGTGGCTGAACA-CTGAGGAGCTGACATCCCATGTC 852
QY 61 AGCAATGTCAATGGCATCAGAGGGGGCCGCCAAGGGCCCATGCG---TTCCCTTCATGC 117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 AGCAATGTCAATGGCATCAGAGGGGGGGCCGCCAAGGGCCCATGCGCTTTCCCTTCATGC 792
QY 118 ATCCATGTGCTGTTTCATTCATTCATCCATACACCTGCTGAGCTTTCACTCT 177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
791 ATCCATGTGCTGTTTCATTCATTCATCCATACACCTGCTGAGCTTTCACTCT 732
QY 178 GACTCCCTAACTCCATCAGACCTCTACGACCATAGACTCTGCCAGAACTGAGAGC 235
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
731 GACTCCCTAACTCCATCAGACCTCTACGACCATAGACTCTGCCAGAACTGAGAGC 674

CB441198 484 bp mRNA linear EST 25-MAR-2003
691539 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
CB441198
CB441198.1 GI:29227980
EST.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 484)
Smith, T.P.L., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G.,
Wray, J.B. and Keeler, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: FQY8048 row: K column: 21
Seq primer: GTAATACGACTCACTATAGG.
FEATURES
Location/Qualifiers
1..484
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."
BASE COUNT 110 a 134 c 107 g 133 t
ORIGIN
Query Match 42.5%; Score 101.2; DB 14; Length 484;
Best Local Similarity 78.5%; Pred. No. 5.9e-18;
Matches 135; Conservative 0; Mismatches 33; Indels 4; Gaps 1;
QY 65 ATGTCAATGGCATCAGAGGGGGCCGCCAAGGGCCCATGCGTTCCCTTCATGCATCCATT 124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4 ATATCACTGATCAGAGGGGTGGCCCTAGGACCATCACTTCCCTTCGCGGCCATC 63
QY 125 GTTCTGTTCAATTCATTCATCCATACATCCACTGCTCTGAGCTTTCACTCTGACTCCC 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 CT----TTCAAGTTTCATTCTCTATACATCCACTCACTCTGAACTTTCACTCTGCTTGC 119

```

```

QY 185 TAACTCCATCAGACCTCTACGACCATCAGACTCTGCCAGAACTGAGAGCC 236
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 TAACTCCACCAAGCTCTACACACCTTTCAGACTTTGCCAGAACTGAGAGCC 171

BX420015 1201 bp mRNA linear EST 13-MAY-2003
BX420015 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF023YA09 5-PRIME, mRNA sequence.
BX420015
BX420015.1 GI:30646738
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6147.r For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF023AA05QP1&cluster=6147.r. Contact:
Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0DF023AA05QPI.
location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 291 a 289 c 334 g 243 t 44 others
ORIGIN
Query Match 41.5%; Score 98.8; DB 13; Length 1201;
Best Local Similarity 67.6%; Pred. No. 4.2e-17;
Matches 127; Conservative 16; Mismatches 44; Indels 1; Gaps 1;
QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGCAACACCTGAGGAGCTGACATCCCATGTC 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
940 AACAGGAGCAGTGTGACATGAGTGGCTGCAACACCTGAGGAGCTGACATCCCATGTC 999
QY 61 AGCAATGTCAATGGCATCAGAGGGGGCCGCCAAGGGCCCATGCG---TTCCCTTCATGC 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1000 ARCAAWTGTAAATTCGATARGAGGGGGCCGCCAAGGG---CCCATCTCTCTTCATGCATC 1058
QY 121 CATGTGTTCTGTTTCATTCATTCATCCATACATCCACTGCTCTGAGCTTTCACTCTGAC 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1059 ATTGTYTTCATTAAWTATACMAAAATCACTKCCCTTKASGTTTACTTGTGTCCTAAW 1118
QY 181 TCCCTAAC 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1119 TCAYTAIV 1126

RESULT 6
AK049286

```

LOCUS AK049286 4149 bp mRNA linear HTC 05-DEC-2002  
 DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library, clone: C330019I03 Product: VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus], full insert sequence.

ACCESSION AK049286  
 VERSION AK049286.1 GI:26093400  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE 1  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 9279253  
 PUBMED 10349636

REFERENCE 2  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE 3  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 JOURNAL sequencing pipeline with 384 multichipillary sequencer  
 MEDLINE Genome Res. 10 (11), 1757-1771 (2000)  
 PUBMED 20530913  
 REFERENCE 11076861

REFERENCE 4  
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staab, J. F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S. and Hayashizaki, Y.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409 (6821), 685-690 (2001)  
 MEDLINE 21085660  
 PUBMED 11217851

REFERENCE 5  
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 REFERENCE 6 (bases 1 to 4149)  
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hara, A., Hayashizaki, Y., Hasegawa, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, F., Hirozane, T.,

TITLE Direct Submission  
 JOURNAL Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/

FEATURES  
 source location/Qualifiers  
 1..4149  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:C330019I03"  
 /db\_xref="taxon:10090"  
 /clone="C330019I03"  
 /cell\_type="ES cells"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 410..1058  
 /note="VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus] (SPTR|Q9JK00, evidence: FASTY, 99.5%ID, 100%length, match=646)  
 putative"  
 4129..4134  
 /note="putative"  
 4149  
 /note="putative"  
 BASE COUNT 1088 a 1022 c 961 g 1077 t 1 others  
 ORIGIN  
 Query Match 17.7%; Score 42.2; DB 11; Length 4149;  
 Best Local Similarity 67.8%; Pred. No. 0.67;  
 Matches 59; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
 QY 36 CTGAGGGACTGGACATCCCATGTTTCAGCAATGTCAATGCGATCGAGGGCGGCCCAAGG 95  
 DB 1310 CTGAGCAACTGGCTATCCCATGTTTCAGCAATGCCAGCAATCATGAGGAAGTCCCCCGAGT 1369  
 QY 96 GCCCCATCGCTCCCTTCATCATCCCA 122  
 DB 1370 GTCCCAACATCATCATCTTTCTATTCA 1396

RESULT 7  
 BG001495/c  
 LOCUS BG001495 322 bp mRNA linear EST 24-JAN-2001  
 DEFINITION MR3-GN02299-141100-003-f07 GN0229 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG001495  
 VERSION BG001495.1 GI:12439883  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 322)

[illegible]

- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

Location/Qualifiers  
1..941  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone\_lib="G"  
/note="Genoscope sequence ID : COAG260BD03LP1-end : T7"

BASE COUNT 203 a 194 c 331 g 203 t 10 others  
ORIGIN

Query Match 17.1%; Score 40.8; DB 29; Length 941;  
Best Local Similarity 51.4%; Pred. No. 0.84;  
Matches 108; Conservative 5; Mismatches 92; Indels 5; Gaps 1;

QY 4 AGAGCAGTGTGACATGAGTGGCTGAACACTGAGGACTGGACATCCCATGTTTCAGC 63  
DB 868 AGTGGCGGGGACAGCTGTTTACGAGAACACCTGTGGAGAGAAACCAAGGTGGGA 809  
QY 64 ATATGTCATGCGATCAGAGGGGCGCCCAAGGCCCCCATCGTTCCTTCATGCAATCCAT 123  
DB 808 CAGCAGAGSGTGAGRGGGGGGACCCGCTCTCCATCCCTCCCTCCATCCATC 752  
QY 124 TGTTCGTTCATTCATTCATCCATACATCCACCTGCTGAGCTTTTACCTCTGACCTCC 183  
DB 751 --ATCCCTCCATCCATCCCTCCATCCATCCCTCCATCCATCCCTCCATCCCTCC 694  
QY 184 CTAATCCATCCAGCTCTACGACCATAA 213  
DB 693 ATCCCTCCATCCATCCCTCCATCCATCCCTCCATCCATCCCTCCATCCCTCC 664

RESULT 10  
CNS02UEH/c  
LOCUS  
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 167623 of library G from Tetraodon nigroviridis, genomic survey sequence.  
AL214370  
ACCESSION AL214370.1 GI:7873189  
VERSION GSS; genome survey sequence.  
KEYWORDS Tetraodon nigroviridis  
SOURCE Tetraodon nigroviridis  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1  
AUTHORS Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.  
TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence  
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)  
MEDLINE 20296633  
PUBMED 10835645

REFERENCE 2  
AUTHORS Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Fzames, C., Bernot, A., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.  
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
JOURNAL Genome Res. 10 (7), 939-949 (2000)  
MEDLINE 20359837  
PUBMED 10899143

REFERENCE 3  
AUTHORS (bases 1 to 844)  
TITLE Direct Submission

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqrefgenoscope.cns.fr](mailto:seqrefgenoscope.cns.fr))  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

Location/Qualifiers  
1..844  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone\_lib="G"  
/note="Genoscope sequence ID : COAG167AD12LP1-end : T7"

BASE COUNT 326 a 118 c 238 g 130 t 32 others  
ORIGIN

Query Match 16.9%; Score 40.2; DB 29; Length 844;  
Best Local Similarity 56.1%; Pred. No. 1.2;  
Matches 60; Conservative 7; Mismatches 40; Indels 0; Gaps 0;

QY 105 CTTCCTTCATGCAATCTGTTCTTCATTCATTCATCCATACATCCATCCCTGCTCTG 164  
DB 661 CATCTTCCATCCATCCATCCATCATATATCYACATCTATCCATCCATCCATCTTT 602  
QY 165 AGCTTTCACCTCTGACCTCCCTPACTCCATCCAGACTCTACGACCAT 211  
DB 601 ATCTCTCCWYCATYCATCTWTCTCATCCATCATATATCTATCCACCT 555

RESULT 11  
BH308136  
LOCUS  
DEFINITION BH230-193A19.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone BH230-193A19, genomic survey sequence.  
BH308136  
ACCESSION BH308136.1 GI:17233605  
VERSION GSS.  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 611)  
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C.M.  
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
JOURNAL Unpublished  
COMMENT Other GSSs: CH230-193A19.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: [szhao@tigr.org](mailto:szhao@tigr.org)  
Clones are derived from the rat BAC library CHORI-230 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library availability, please contact Pieter de Jong ([pdj@tigr.org](mailto:pdj@tigr.org)). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/oreringinformation.htm>). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html)  
Plate: 193 row: A column: 19  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers  
1..611  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SeNHsd/MCW"  
/db\_xref="taxon:10116"

```

/clone="CH230-193A19"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/notes="Vector: PTABAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
Pieter de Jong"
BASE COUNT      182 a   144 c   74 g   211 t
ORIGIN

Query Match      16.8%; Score 40; DB 28; Length 611;
Best Local Similarity 58.3%; Pred. NO. 1.2;
Matches 70; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 105 CTTCCCTTCATGCATCCATTGTTCTGTTTCATTCATTCATCCATCCATCCACCTGCTGTC 164
Db 432 CATCCATCATCATCATCCATCCCTTCCTGTCGTCATCCATCCCTCGCTCGTCTGTCGTC 491
QY 165 AGCTTTACCTCTGACTCCCTAACTCCATCAGACCTCTACGACCAATAGACTCTGCCAG 224
Db 492 TGTCTGTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 551

RESULT 12
AQ835006
LOCUS
DEFINITION
HS 3173_B2_H08_T7A CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plates=3173 Col=16 Row=P, Genomic survey
sequence.
ACCESSION
AQ835006
VERSION
AQ835006.1 GI:5801068
SOURCE
GSS.
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 676)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end web server: http://www.htsc.washington.edu
Plate: 3173 row: P column: 16
Seq primer: T7
Class: BAC ends
High quality sequence stop: 676.
Location/Qualifiers
1..676
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3173 Col=16 Row=P"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/notes="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      163 a   209 c   72 g   211 t   21 others
ORIGIN

Query Match      16.5%; Score 39.2; DB 28; Length 676;
Best Local Similarity 60.8%; Pred. No. 2;

```

```

Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 105 CTTCCCTTCATGCATCCATTGTTCTGTTTCATTCATTCATCCATCCATCCACCTGCTGTC 164
Db 329 CATCCATCATCATCATCCATCCATCCATTCATTCATTCATCCATCCATCCATCCATCCATCC 388
QY 165 AGCTTTACCTCTGACTCCCTAACTCCATCAGACCTCTACGACCAATAGACTCTGCCAG 201
Db 389 ATCAGTCATCATCTATAGGTCATCATCATCATCTC 425

RESULT 13
AZ627327/c
LOCUS
DEFINITION
1M0469M04F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0469M04 F, genomic survey sequence.
ACCESSION
AZ627327
VERSION
AZ627327.1 GI:11749517
SOURCE
GSS.
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 564)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenan,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weise,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0469 row: M column: 04
Seq primer: CGTTGTAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 564.
Location/Qualifiers
1..564
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0469M04"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/notes="vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
FEATURES
source

```

```

BASE COUNT      174 a   108 c   150 g   132 t
ORIGIN
Query Match      16.4%; Score 39; DB 28; Length 564;
Best Local Similarity 61.2%; Pred. No. 2.1;
Matches 63; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 100 CATCGCTCCCTTCATCCATCCATTCCTTCCTTCATTCATTCATCCATCCATCCCTGC 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 CATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 CTCTGAGCTTCACCTCTGACTCCCTAACTCCATCCATCCAGACCTCT 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 ACTGCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
BY730622
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BY730622 660 bp mRNA linear EST 17-DEC-2002
musculus cDNA clone D530025M20 5', mRNA sequence.
BY730622
BY730622.1 GI:27143749
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 660)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedierski, R. M.,
King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, B., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomata, M., Verardo, R., Wagner, L., Wanstedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Atakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
1246851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome.res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno,
H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,

```

```

Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,
Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1..660
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D530025M20"
/tissue_type="stomach"
/dev_stage="13 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 13 days embryo
stomach"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGCGCCGACCTCGAGTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGATCTCGAGTTTAAATTAATCCCTCCCTCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 124 a 171 c 134 g 226 t
ORIGIN
Query Match 16.2%; Score 38.6; DB 14; Length 660;
Best Local Similarity 68.8%; Pred. No. 3;
Matches 53; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 105 CTTCCTTCATCCATTCATTCCTTCCTTCATTCATTCATCCATCCATCCCTGCTG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 CATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCTGC 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 AGCTTTACCTCTGACT 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 TGCITTCATCTGCT 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
CNS05122/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

CNS05122 1101 bp DNA linear GSS 01-SEP-2000
Tetradon nigroviridis genome survey sequence T7 and of clone
005D19 of library A from Tetradon nigroviridis, genomic survey
sequence.
AL338372
AL338372.1 GI:8232130
GSS; genome survey sequence.

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:39:02 ; Search time 99.088 Seconds

(without alignments)  
7982.969 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_1024\_1261

Perfect score: 238

Sequence: 1 aacaggagcagtgacatg.....tgccagaactgagaagccgg 238

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubnpa/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubnpa/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubnpa/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubnpa/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubnpa/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubnpa/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubnpa/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubnpa/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubnpa/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubnpa/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubnpa/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubnpa/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubnpa/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubnpa/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubnpa/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubnpa/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubnpa/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	238	100.0	1261	10	US-09-997-579-4
2	216	90.8	4625	11	Sequence 4, Appli
3	38	16.0	27079	12	Sequence 7659, Ap
4	38	16.0	38360	11	Sequence 55, Appl
5	37.8	15.9	1091	12	Sequence 14, Appl
6	37.8	15.9	1091	12	Sequence 30772, A
7	37.8	15.9	1091	12	Sequence 30772, A
8	37.8	15.9	1327	13	Sequence 215913,
9	37.8	15.9	3768	10	Sequence 215913,
10	37.8	15.9	3768	10	Sequence 708-464
11	37.8	15.9	3768	10	Sequence 464, App
12	37.8	15.9	3768	12	Sequence 164, App
13	37.8	15.9	3768	12	Sequence 373, App
14	37.6	15.8	1781	12	Sequence 689, App
15	37.6	15.8	1781	12	Sequence 1041, Ap
16	37	15.5	2632	13	Sequence 200093,
					Sequence 22, Appl

17	37	15.5	3108	13	US-10-029-191-1	Sequence 1, Appli
18	37	15.5	6419	12	US-10-311-455-239	Sequence 239, App
19	36.6	15.4	494	12	US-10-027-632-320641	Sequence 320641,
20	36.6	15.4	494	13	US-10-027-632-320641	Sequence 320641,
21	36.6	15.4	642	12	US-10-027-632-61529	Sequence 61529, A
22	36.6	15.4	642	13	US-10-027-632-61529	Sequence 61529, A
23	36.6	15.4	4743	11	US-09-764-891-7898	Sequence 7898, Ap
24	36.6	15.4	4743	11	US-09-764-891-7897	Sequence 7897, Ap
25	36.4	15.3	29322	12	US-10-034-650-58	Sequence 58, Appl
26	36.2	15.2	615	12	US-10-027-632-278139	Sequence 278139,
27	36.2	15.2	615	13	US-10-027-632-278139	Sequence 278139,
28	36.2	15.2	626	12	US-10-027-632-278138	Sequence 278138,
29	36.2	15.2	626	13	US-10-027-632-278138	Sequence 278138,
30	36.2	15.2	805	12	US-10-027-632-169441	Sequence 169441,
31	36.2	15.2	805	13	US-10-027-632-169441	Sequence 169441,
32	36	15.1	473	12	US-10-027-632-41251	Sequence 41251, A
33	36	15.1	473	13	US-10-027-632-41251	Sequence 41251, A
34	36	15.1	612	12	US-10-027-632-22189	Sequence 22189, A
35	36	15.1	612	13	US-10-027-632-22189	Sequence 22189, A
36	36	15.1	685	12	US-10-027-632-147227	Sequence 147227,
37	36	15.1	685	12	US-10-027-632-147228	Sequence 147228,
38	36	15.1	685	12	US-10-027-632-147229	Sequence 147229,
39	36	15.1	685	13	US-10-027-632-147227	Sequence 147227,
40	36	15.1	685	13	US-10-027-632-147228	Sequence 147228,
41	36	15.1	685	13	US-10-027-632-147229	Sequence 147229,
42	35.8	15.0	739	12	US-10-027-632-135770	Sequence 135770,
43	35.8	15.0	739	12	US-10-027-632-135771	Sequence 135771,
44	35.8	15.0	739	13	US-10-027-632-135770	Sequence 135770,
45	35.8	15.0	739	13	US-10-027-632-135771	Sequence 135771,

ALIGNMENTS

RESULT 1

US-09-997-579-4

; Sequence 4, Application US/09997579

; Patent No. US20020113203A1

; GENERAL INFORMATION:

; APPLICANT: Cambridge University Technical Services

; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated s

; TITLE OF INVENTION: channel

; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses

; FILE REFERENCES: 674558-2001

; CURRENT APPLICATION NUMBER: US/09/997,579

; CURRENT FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: PCT/EP00/01783

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60,129,473

; PRIOR FILING DATE: 2000-02-24

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1261

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-997-579-4

Query Match 100.0%; Score 238; DB 10; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AACAGGACGAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGACATCCCATGTC 60
Db	1024	AACAGGACGAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGACATCCCATGTC 1083
Qy	61	AGCAATGTCATGAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGACATCCCATGTC 120
Db	1084	AGCAATGTCATGAGTGTGACATGAGTGGCTGACACCTGAGGAGCTGACATCCCATGTC 1143
Qy	121	CATTGCTGCTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 180
Db	1144	CATTGCTGCTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1203

**Qy**      181    TCCCTAACTCCATCAGACCTCTACGCACCATAAGACTCTGCCAGAAGCGG    238  
             |||||  
**Dδ**      1204    TCCCTAACTCCATCAGACCTCTACGCACCATAAGACTCTGCCAGAAGCGG    1261  
             |||||

## RESULT 2

```

US-09-764-891-7659
; Sequence 7659, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7659
; LENGTH: 4625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7659

```

### RESULT, T 3

```

RESULTS 3
US-10-034-650-55
; Sequence 55, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1993-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 27079
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-034-650-55

```

## RESULT, T 4

```

US-09-999-121-14
; Sequence 14, Application US/09999121
; Publication No. US20030039982A1
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; TITLE OF INVENTION: ISOLATED GENOMIC NUCLEOTIDE FRAGMENTS FROM THE p15 REGION OF
; TITLE OF INVENTION: CHROMOSOME 11
; FILE REFERENCE: JR-10,001-US
; CURRENT APPLICATION NUMBER: US/09/999,121
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US 60/244,705
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 38360
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-999-121-14

```

## RESULTS

```

RESULTS 5
US-10-027-632-30772/c
; Sequence 30772, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30772
; LENGTH: 1091

```







Qy 167 CTTTCACCTCTGACTC 182  
| | | | |  
Db 161 TCTACACATATTATC 176

Search completed: December 5, 2003, 23:17:03  
Job time : 100.088 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 23:22:39 ; Search time 4840 Seconds  
(without alignments)  
10658.479 Million cell updates/sec

Title: US-09-936-680-4  
Perfect score: 1261  
Sequence: 1 cctccctccgagctgagc.....tgccagaactgagaagccgg 1261

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size : 10

Total number of hits satisfying chosen parameters: 2382734

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pin.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sv.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1261	100.0	1261	6	AX039100	Sequence
2	1261	100.0	1261	9	HS2423396	AX039100 Homo sapi
3	1104	87.5	4052	9	HS2423396	AL136589 Homo sapi
4	860	68.2	5306	9	AB032984	AP000682 Homo sapi
5	428	33.9	127347	2	AP000682	AC063921 Homo sapi
6	428	33.9	144833	2	AC063921	AC021981 Homo sapi
7	428	33.9	149800	2	AC021981	AC024604 Homo sapi
8	428	33.9	172546	2	AC024604	AP002765 Homo sapi
9	428	33.9	178169	9	AP002765	AC069539 Homo sapi
10	428	33.9	181471	9	AC069539	BD059018 Secreted
11	324	25.7	471	6	BD059018	AP002749 Homo sapi
12	262	20.8	178431	2	AP002749	AC024604 Homo sapi
13	226	17.9	172546	2	AC024604	AB097521 Macaca fa
14	220	17.4	3296	9	AB097521	AC063921 Homo sapi
15	218	17.3	144833	2	AC063921	AC021981 Homo sapi
16	218	17.3	149800	2	AC021981	AP002749 Homo sapi
17	216	17.1	178431	2	AP002749	AC013796 Homo sapi
18	143	11.3	66980	2	AC013796	AC013796 Homo sapi
19	129	10.2	66980	2	AC013796	AX048004 Sequence
20	38	3.0	645	6	AX048004	AX039099 Sequence
21	38	3.0	2220	6	AX039099	AJ243395 Rattus no
22	38	3.0	2220	10	RNO243395	AX048005 Sequence
23	38	3.0	2632	6	AX048005	AF378093 Rattus no
24	38	3.0	3107	10	AF378093	AX047984 Sequence
25	38	3.0	3108	6	AX047984	AY049036 Mus muscu
26	29	2.3	670	10	AY049036	AC135353 Mus muscu
27	29	2.3	143624	2	AC135353	AC128723 Rattus no
28	26	2.1	214306	2	AC128723	AC079118 Homo sapi
29	24	1.9	158801	9	AC079118	AY009586 Boophilus
30	23	1.8	257	3	AY009586	AL645845 Mouse DNA
31	23	1.8	109446	10	AL645845	AC017063 Homo sapi
32	23	1.8	193092	9	AC017063	AC113159 Rattus no
33	23	1.8	196383	2	AC113159	AC005528 Mus muscu
34	23	1.8	198862	10	AC005528	AC127850 Rattus no
35	23	1.8	245938	2	AC127850	AC119520 Rattus no
36	23	1.8	254497	2	AC119520	AF102248 Odocoileu
37	23	1.8	271454	2	AC112098	AC135108 Mus muscu
38	22	1.7	518	4	AF102248	AL161448 Human DNA
39	22	1.7	23232	2	AC135108	EX537251 Danio rer
40	22	1.7	164857	9	AL161448	AL627256 Zebrafish
41	22	1.7	165095	2	EX537251	EX088532 Danio rer
42	22	1.7	169757	5	AL627256	AC011242 Homo sapi
43	22	1.7	175930	2	EX088532	AC114645 Mus muscu
44	22	1.7	186091	9	AC011242	
45	22	1.7	198896	2	AC114645	

# ALIGNMENTS

RESULT 1  
AX039100  
LOCUS AX039100  
DEFINITION Sequence 4 from Patent WO0063367.  
ACCESSION AX039100  
VERSION AX039100.1 GI:11229276  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Cox, P., Dixon, A., Jackson, A. and Morgan, K.  
TITLE A novel family of beta sub-unit proteins from a voltage-gated sodi  
um channel, nucleic acids encoding them and therapeutic or

1  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
linear PAT 18-NOV-2000





```

DH10B; sites NotI + SalI
/dev stage="adult"
1. 4052
/gene="DKF2p761f182"
804..1451
/gene="DKF2p761f182"
/notes="scn3b (Homo sapiens)"
/codon start=1
/product="hypothetical protein"
/protein_id="CAB6524.1"
/db_xref="GI:13276681"
/db_xref="SWISS-PROT:Q9NY72"
/translation="MPAFNRLFLSLVLIYVVCVFPVCEVPEVETAEVQNPMLR
CISCMKEVEATVYVWFREYREGKDFLIYRNHGOEVSFPQRIOWNGSKDLQD
VSTLVNVLNDSGLYCNVSRFEFAHRPFVKTRLIPLVTEAEGEDFTSVVSEI
MVIYLLVFLTLWLLIEMVYCYRKVSKABEAAQENASDYLAIPESENKENSAPVBE"
polyA_site
4035
/gene="DKF2p761f182"
BASE COUNT 994 a 1033 c 1036 g 989 t
ORIGIN
Query Match 87.5%; Score 1104; DB 9; Length 4052;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 CTCCTTCGAGCTGAGCTTACCTGGGCGCAACAGCGAGGCGAGGCGGCGAGTGGAA 62
Db 431 CTCCTTCGAGCTGAGCTTACCTGGGCGCAACAGCGAGGCGAGGCGGCGAGTGGAA 490
Qy 63 GCTGGAGTTCGGGCTGGGCGGAGGCGGAGCTGTCGCTGGTGTGAGCGCGGCGAGAGC 122
Db 491 GCTGGAGTTCGGGCTGGGCGGAGGCGGAGCTGTCGCTGGTGTGAGCGCGGCGAGAGC 550
Qy 123 GGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTGCCTTAGG 182
Db 551 GGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTGCCTTAGG 610
Qy 183 CCCAAGCCCCACCGCGCTCCAAAGCTCCAGGGGCTCCCGAGGCGACCGGTGCTCGGC 242
Db 611 CCCAAGCCCCACCGCGCTCCAAAGCTCCAGGGGCTCCCGAGGCGACCGGTGCTCGGC 670
Qy 243 CTTCTTCGCTCAGAAAGTGGCCCCCTGGGGGAGTGTGCTCCAAAGGTTTCCTCGAA 302
Db 671 CTTCTTCGCTCAGAAAGTGGCCCCCTGGGGGAGTGTGCTCCAAAGGTTTCCTCGAA 730
Qy 303 AGAATCTGAGAGGCGGAGTCTGTGACGAGGGAATCTCTGTGTAGCTTGGAAGCG 362
Db 731 AGAATCTGAGAGGCGGAGTCTGTGACGAGGGAATCTCTGTGTAGCTTGGAAGCG 790
Qy 363 CCAGCCCCCAGAAAGTGGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTAT 422
Db 791 CCAGCCCCCAGAAAGTGGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTAT 850
Qy 423 CTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 482
Db 851 CTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 910
Qy 483 GCAGGCAACCCCATGAAGTGTGCTGCTTCTTCCATGAGAGAGAGAGAGAGAGAG 542
Db 911 GCAGGCAACCCCATGAAGTGTGCTGCTTCTTCCATGAGAGAGAGAGAGAGAGAG 970
Qy 543 CACACCGGTGTGAAGTGTGTCTACAGGCGGAGGCGGTAAAGTTCCTTATTTACGA 602
Db 971 CACACCGGTGTGAAGTGTGTCTACAGGCGGAGGCGGTAAAGTTCCTTATTTACGA 1030
Qy 603 GTATCGGAATGCCACAGGAGGTGGAGAGCCCCCTTTCAGGGGCGCTCGAGTGGATGG 662
Db 1031 GTATCGGAATGCCACAGGAGGTGGAGAGCCCCCTTTCAGGGGCGCTCGAGTGGATGG 1090
Qy 663 CAGCAAGACCTCGCAGAGAGTGTCCATCTGCTCAAGTCAAGTCAAGTCAAGTCAAG 722
Db 1091 CAGCAAGACCTCGCAGGAGCTGTCTCATCTGTGCTCAAGTCAAGTCAAGTCAAG 1150

723 CCTCTACACCTGCAATGTGTCCCGGAGTTTGTAGTTTGAGGCGCATCGGCCCTTTGTGAA 782
1151 CCTCTACACCTGCAATGTGTCCCGGAGTTTGTAGTTTGAGGCGCATCGGCCCTTTGTGAA 1210
783 GACGACGGGCTGATCCCTTAAGAGTCAACGAGGAGGCTGGAGAGGACTTACACCTCTGT 842
1211 GACGACGGGCTGATCCCTTAAGAGTCAACGAGGAGGCTGGAGAGGACTTACACCTCTGT 1270
843 GGTCTCAGAAATCATGATGATACATCTTCTGTGTCTTCTCACCCTGTGGCTGTCTATCGA 902
1271 GGTCTCAGAAATCATGATGATACATCTTCTGTGTCTTCTCACCCTGTGGCTGTCTATCGA 1330
903 GATGATATATGTACAGAAAGTCTCAAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 962
1331 GATGATATATGTACAGAAAGTCTCAAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1390
963 TGACTACTCTTGCCATCCATCTGAGAAACAGGAGAACTCTGCGGTACCACTGGAGGAATA 1022
1391 TGACTACTCTTGCCATCCATCTGAGAAACAGGAGAACTCTGCGGTACCACTGGAGGAATA 1450
1023 GAACAGGAGGAGTGTGATGAGTGGGCTTGAAACACTGAGGAGGAGTGGAGCATCCCATGTT 1082
1451 GAACAGGAGGAGTGTGATGAGTGGGCTTGAAACACTGAGGAGGAGTGGAGCATCCCATGTT 1510
1083 CAGCAATGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATGCTTCCCTTCATGCAT 1142
1511 CAGCAATGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATGCTTCCCTTCATGCAT 1570
1143 CCATGTGTCTGTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1202
1571 CCATGTGTCTGTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1630
1203 CTCCTACTCCATCAGTACGCTCTACGACCATGAGTCTGCGGAGGAGTGGAGAGGC 1259
1631 CTCCTACTCCATCAGTACGCTCTACGACCATGAGTCTGCGGAGGAGTGGAGAGGC 1687

RESULT 4
AB032984
LOCUS Homo sapiens mRNA for KIAA1158 protein, partial cds.
DEFINITION Homo sapiens mRNA for KIAA1158 protein, partial cds.
ACCESSION AB032984
VERSION AB032984.1 GI:6330135
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Hiroseawa,M., Nagase,T., Ishikawa,K., Kikuno,R., Nomura,N. and
Ohara,O.
TITLE Characterization of cDNA clones selected by the GenesMark analysis
from size-fractionated cDNA libraries from human brain
JOURNAL DNA Res. 6 (5), 329-336 (1999)
MEDLINE 20039618
PubMed 10574461
REFERENCE 2 (bases 1 to 5306)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1512-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp, Tel:+81-438-52-3913,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3914,
Fax:+81-438-52-3914)
FEATURES
Location/Qualifiers
source
1..5306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hJ00081"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
```

```

/dev_stage="adult"
1..5306
/gene="KIAA1158"
<1..695
/gene="KIAA1158"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1158 protein"
/protein_id="BAA86472.1"
/db_xref="GI:6330136"
/translation="RSWGSGESLGRQPMFAPNRLFLPLASLVLYWVSVPFVCEV
PSETEAVQGNPKRLRCISCKREVEEATVWVFVPGCKDFLYEVRNGHOEYBSP
RGRLOMNSKDLQDVSIIVLVNLTNDLSGLTCNVSREFEARHPFKVTRLLPLRV
TERAGDFTSVSEINMYILLVFLILWLLIEMCYRKVSKAEBAQAENASDYLAIPS
ENKENSAPVEE"
BASE COUNT 1375 a 1311 c 1111 g 1509 t
ORIGIN
Query Match 68.2%; Score 860; DB 9; Length 5306;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 349 AGCTTGGAGAGCGCCAGAGAGATGCTTCAATAGATGTTTCCCTGGCT 408
DB 21 AGCTTGGAGAGCGCCAGAGAGATGCTTCAATAGATGTTTCCCTGGCT 80
QY 409 TCCTCTGCTTATCTACTGGTCACTGCTGCTTCCCTGCTGTTGTTGGAAGTCCCTCG 468
DB 81 TCCTCTGCTTATCTACTGGTCACTGCTGCTTCCCTGCTGTTGTTGGAAGTCCCTCG 140
QY 469 GAGACGAGGCGCTGAGGAGCAACCCATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTG 528
DB 141 GAGACGAGGCGCTGAGGAGCAACCCATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTG 200
QY 529 GAGAGGTGAGGCGCCAGGAGGAGTGGTCTTACAGGCGGAGGCGGTAAGAT 588
DB 201 GAGAGGTGAGGCGCCAGGAGGAGTGGTCTTACAGGCGGAGGCGGTAAGAT 260
QY 589 TTCTTTATTTACGATGATCGAATGGCCACAGAGGAGTGGAGAGGCGGCTTTCAGGGGCGC 648
DB 261 TTCTTTATTTACGATGATCGAATGGCCACAGAGGAGTGGAGAGGCGGCTTTCAGGGGCGC 320
QY 649 CTGAGTGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 708
DB 321 CTGAGTGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 380
QY 709 CTGAACGACCTCTGGCTCTACACCTGCAATGTTGCTCCGGAGGTTTGAGTTTGAAGCGCAT 768
DB 381 CTGAACGACCTCTGGCTCTACACCTGCAATGTTGCTCCGGAGGTTTGAGTTTGAAGCGCAT 440
QY 769 CGGCGCTTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 828
DB 441 CGGCGCTTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 500
QY 829 GACTTACCTCTCTGTGCTCTCAGAAATCATGATGATGATGATGATGATGATGATGATGATG 888
DB 501 GACTTACCTCTCTGTGCTCTCAGAAATCATGATGATGATGATGATGATGATGATGATGATG 560
QY 889 TGGCTGCTATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 948
DB 561 TGGCTGCTATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 620
QY 949 CAGAAAGGCGCTGACTACCTTGCATCCCATCTGAGAACAGAGGAGGAGGAGGAGGAGGAGG 1008
DB 621 CAGAAAGGCGCTGACTACCTTGCATCCCATCTGAGAACAGAGGAGGAGGAGGAGGAGGAG 680
QY 1009 CCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1068
DB 681 CCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 740
QY 1069 GGCATCCCATGCTGACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1128
DB 741 GGCATCCCATGCTGACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 800

```

```

QY 1129 TTCCCTTCATGATCATTTGTTCTGTTCAATTCATTCATTCATTCATTCATTCATTCATTCAT 1188
DB 801 TTCCCTTCATGATCATTTGTTCTGTTCAATTCATTCATTCATTCATTCATTCATTCATTCAT 860
QY 1189 GCTTTACCTCTGACTCCCTAACTCCATCAGACCTCTAGCCACCAAGACTCTGCAGAG 1248
DB 861 GCTTTACCTCTGACTCCCTAACTCCATCAGACCTCTAGCCACCAAGACTCTGCAGAG 920
QY 1249 ACTGAGAAGCC 1259
DB 921 ACTGAGAAGCC 931

```

## RESULT 5

AP000682/c

LOCUS

DEFINITION Homo sapiens chromosome 11 clone CMB9-32A1 map 11q24, WORKING DRAFT  
SEQUENCE, 8 unordered pieces.

ACCESSION

AP000682

VERSION

AP000682.3 GI:9844967

KEYWORDS

HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 127347)

AUTHORS

Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.

TITLE

Homo sapiens 127,347 genomic DNA of 11q24

REFERENCE

2 (bases 1 to 127347)

AUTHORS

Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.

TITLE

Direct Submission

JOURNAL

Submitted (08-NOV-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-15-1 Kitagato, Sagamihara, Kanagawa 228-8555, Japan

COMMENT

Tel: 81-42-778-9923, Fax: 81-42-778-9924

On Aug 19, 2000 this sequence version replaced gi:8118870.

----- Genome Center

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: HumDraft11

Center clone name: CMB9-32A1

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 124981 bases at least Q40

Consensus quality: 125815 bases at least Q30

Consensus quality: 126286 bases at least Q20

Insert size: 126647; sum-of-contigs

Quality coverage: 12.71x in Q20 bases; sum-of-contigs

-----

NOTE: This is a 'working draft' sequence. It currently consists of

8 contigs. The true order of the pieces is not known and their

order in this sequence record is arbitrary. Gaps between the

contigs are represented as runs N, but the exact sizes of the gaps

are unknown. This record will be updated with the finished sequence

as soon as it is available and the accession number will be

preserved

1

33532 contig of 33532 bp in length

33633

59223 contig of 25591 bp in length

59324

83007 contig of 23684 bp in length

83108

100196 contig of 17089 bp in length

100297

112313 contig of 12017 bp in length

112414

121043 contig of 8630 bp in length

121144

125917 contig of 4774 bp in length

126018 127347 contig of 1330 bp in length.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

*	1	33532: contig of 33532 bp in length
*	33533	33632: gap of 100 bp
*	33633	59223: contig of 25591 bp in length
*	59224	59323: gap of 100 bp
*	59324	83007: contig of 23684 bp in length
*	83008	83107: gap of 100 bp
*	83108	100136: contig of 17089 bp in length
*	100137	100296: gap of 100 bp
*	100297	11233: contig of 12017 bp in length
*	11234	112413: gap of 100 bp
*	112414	121043: contig of 8630 bp in length
*	121044	121143: gap of 100 bp
*	121144	125917: contig of 4774 bp in length
*	125918	126017: gap of 100 bp
*	126018	127347: contig of 1330 bp in length.

## FEATURES

```
1. 127347
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q24"
/clone="CMB9-32A1"
```

BASE COUNT	36325 a	27508 c	27825 g	34989 t	700 others
ORIGIN					
Query Match	33.9%	Score 428;	DB 2;	Length 127347;	
Best Local Similarity	100.0%;	Pred. No. 1.2e-233;			
Matches 428;	Conservative 0;	Mismatches -0;	Indels 0;	Gaps 0;	
QY	3	CTCCCTCCGAGCTGAGCTTAACCTCGGGGCGCAAAACAGCGAGCGCGAGCGCGCGAGTCGAA	62		
Db	29132	CTCCCTTCGAGCTGAGCTTAACCTCGGGCGCAAAACGAGCGAGCGCGAGCGCGAGTCGAA	29073		
QY	63	GCTGGAGTTCCGGGCTGGGCGGGGAGCGGACGTGCTCGTGCTGCTGAGCGCCGGCGGAGAC	122		
Db	29072	GCTGGAGTTCCGGGCTGGGCGGGGAGCGGACGTGCTCGTGCTGCTGAGCGCCGGCGGAGAC	29013		
QY	123	GGCGCGGAGCGGCTCATCGGCTCCCTCGAACTGGGGAGGTCACAGTGGGGTCGCTTAGGG	182		
Db	29012	GGCGCGGAGCGGCTCATCGGCTCCCTCGAACTGGGGAGGTCACAGTGGGGTCGCTTAGGG	28953		
QY	183	CCCAAGCCGCCACCCGGCTCCAAAGCTCCAGGCGCTCCCGAGGCAACCGGTGCTCGGC	242		
Db	28952	CCCAAGCCGCCACCCGGCTCCAAAGCTCCCGAGGCGCTCCCGAGGCAACCGGTGCTCGGC	28893		
QY	243	CTTTCCTTCGGTCAGAAAGTCGCCCCCTGGGGGCGAGTTCGTCCTCCCAAGGGTTTCTCGAA	302		

28992 CCTTCTTCGGT CAGAAAGTCGCCCTCGGGGGCAGTTCGTGCCCAAGGGTTCTTCGAA 28993

303 AGAATCTGAGAGGGCGCAGTCTCTTGACCGAGGGAATCTCTCTGTGTAGCTTGGGAAGCG 362

28932 AGAATCTGAGAGGGCGCAGTCTCTTGACCGAGGGAATCTCTCTGTGTAGCTTGGGAAGCG 28773

363 CCGAGCCCGCAGAGATGCGCTTGTCTTAATAGATGTTTCCCTGGCTTCTCTCGTGCCTTAT 422

28772 CCGAGCCCGCAGAGATGCGCTTGTCTTAATAGATGTTTCCCTGGCTTCTCTCGTGCCTTAT 28713

423 CTACTGGG 430

28712 CTACTGGG 28705

RESULT 6  
AC063921/c  
LOCUS  
DEFINITION  
AC063921  
AC063921  
HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 144833)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbiera,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Dunthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Halle,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,B., Liu,W., Loulseg,H.,  
Lopez,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mahoney,E., McLeod,M., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,X., Morgan,M., Morris,N.,  
Mosser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogih,M., Okwuonu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,J.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojblomak,I., Rolfe,M., Ruiz,S., Savory,G.,  
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sison,I.,  
Sodergren,E., Sone,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Williams,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,X.,  
Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
Direct Submission  
Unpublished  
2 (bases 1 to 144833)  
Worley,K.C.  
Direct Submission

## JOURNAL

Submitted (22-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

3 (bases 1 to 144833)  
Worley, K.C.  
Direct Submission  
Submitted (08-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Apr 28, 2002 this sequence version replaced gi:16117924.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Project name: HAXY  
Center project name: HAXY  
Center clone name: RP11-142P10  
----- Summary Statistics

Sequencing vector: M13;  
Chemistry: Dye-terminator Big Dye; 14% of reads  
Assembly: Dye-terminator Big Dye; 86% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 127260 bases at least Q40  
Consensus quality: 132735 bases at least Q30  
Consensus quality: 136183 bases at least Q20  
Estimated insert size: 139078; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 13 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2437: contig of 2437 bp in length  
2438 2537: gap of unknown length  
2538 4812: contig of 2275 bp in length  
4813 4913: gap of unknown length  
4913 8488: contig of 3576 bp in length  
8488 8588: gap of unknown length  
8588 11163: contig of 2575 bp in length  
11163 11263: gap of unknown length  
11263 16638: contig of 5375 bp in length  
16638 16738: gap of unknown length  
16738 21870: contig of 5132 bp in length  
21870 21970: gap of unknown length  
21970 30780: contig of 8810 bp in length  
30780 30880: gap of unknown length  
30880 39267: contig of 8387 bp in length  
39267 39368: gap of unknown length  
39368 46890: contig of 7523 bp in length  
46890 46990: gap of unknown length  
46990 59656: contig of 12666 bp in length  
59656 59756: gap of unknown length  
59756 83309: contig of 23553 bp in length  
83309 83409: gap of unknown length  
83409 111857: contig of 28448 bp in length  
111857 111957: gap of unknown length  
111957 144833: contig of 32876 bp in length.

## FEATURES

## source

1. 144833  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-142P10"

## BASE COUNT

## ORIGIN

38456 a 33920 c 34115 g 37093 t 1249 others

Query Match 33.9%; Score 428; DB 2; Length 144833;  
Best Local Similarity 100.0%; Pred. NO. 1.2e-233;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 CTCCTTCGAGCTGAGCTTACCTCGCGCGCAAAACGAGGAGGCGGCGGAGTGGAA 62  
Dy 50353 CTCCTTCGAGCTGAGCTTACCTCGCGCGCAAAACGAGGAGGCGGCGGAGTGGAA 50294  
Qy 63 GCTGAGGTTCCGGGCGGCGGAGGCGGAGTCTCGTGGTGTGAGCGCGCGGAGAGC 122  
Dy 50293 GCTGAGGTTCCGGGCGGCGGAGGCGGAGTCTCGTGGTGTGAGCGCGCGGAGAGC 50234  
Qy 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGCTCCAGTGGGCTCGCTTAGG 182  
Dy 50233 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGCTCCAGTGGGCTCGCTTAGG 50174  
Qy 183 CCCAAAGCCCCCACCGGCTCCAAAGCTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCCAAG 242  
Dy 50173 CCCAAAGCCCCCACCGGCTCCAAAGCTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCCAAG 50114  
Qy 243 CTTCTCTGCTGAGGAGTGGCTGGGCGGCGGCTGGGCGGCGGCTGGGCGGCGGCTGGGCGG 302  
Dy 50113 CTTCTCTGCTGAGGAGTGGCTGGGCGGCGGCTGGGCGGCGGCTGGGCGGCGGCTGGGCGG 50054  
Qy 303 AGAATCTGAGAGGCGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGGCTTTGGAGCGG 362  
Dy 50053 AGAATCTGAGAGGCGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGGCTTTGGAGCGG 49994  
Qy 363 CCAGCCCCCAGAGATGCTGCTTCAATAGATGTTTCCCTGGCTTCTCTGCTGCTTAT 422  
Dy 49993 CCAGCCCCCAGAGATGCTGCTTCAATAGATGTTTCCCTGGCTTCTCTGCTGCTTAT 49934  
Qy 423 CTACTGGG 430  
Dy 49933 CTACTGGG 49926

## RESULT 7

## AC021981/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

149800 bp DNA linear HTG 19-JUL-2000  
Homo sapiens chromosome 11 clone RP11-11C15 map 11, WORKING DRAFT  
SEQUENCE, 29 unordered pieces.  
AC021981.2 GI:9280747  
HTG; HTGS PHASE1; HTGS\_DRAFT.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 149800)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Alderton, J.,  
Anderson, S., Baldwin, J., Barna, N., Becker, R., Bieda, F.,  
Boguslavsky, L., Bouckgeer, B., Brown, A., Burkett, G., Castle, A.,  
Choe, Y., Collange, M., Collins, S., Collins, S., Collins, S.,  
DeArnell, K., Dewar, K., Domino, M., Doyle, M., Fenebor, J.,  
Ferrel, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardner, S., Grant, G., Hargis, B., Heaford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,  
Landers, T., Lech, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
MacDonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPherson, R., Meldrum, J., Menees, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,  
Pierre, N., Pisani, C., Pollack, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Tirrell, A., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Zimmer, A., and Zody, M.  
Direct Submission

## JOURNAL

Submitted (23-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Jul 19, 2000 this sequence version replaced g1:6731265.  
All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1335

Center clone name: L1 C.15

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 135317 bases at least Q40

Consensus quality: 141872 bases at least Q30

Consensus quality: 144404 bases at least Q20

Insert size: 121000; agarose-fp

Quality coverage: 4.2 in Q20 bases; agarose-fp

Quality coverage: 3.5 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 29 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1 1441: contig of 1441 bp in length
1442 1541: gap of 100 bp
1542 2883: contig of 1342 bp in length
2884 2984: gap of 100 bp
2984 4649: contig of 1666 bp in length
4650 4750: gap of 100 bp
4750 6112: contig of 1363 bp in length
6113 7413: contig of 1201 bp in length
7414 7514: gap of 100 bp
7514 9051: contig of 1537 bp in length
9051 10944: contig of 1794 bp in length
10945 11045: gap of 100 bp
11045 13203: contig of 2159 bp in length
13204 13304: gap of 100 bp
13304 16119: contig of 2815 bp in length
16119 16219: gap of 100 bp
16219 18673: contig of 2455 bp in length
18674 18774: gap of 100 bp
18774 21180: contig of 2407 bp in length
21181 21281: gap of 100 bp
21281 23200: contig of 1920 bp in length
23201 23301: gap of 100 bp
23301 26526: contig of 3226 bp in length
26527 29655: contig of 3029 bp in length
29656 34738: contig of 4983 bp in length
34739 34839: gap of 100 bp
34839 39394: contig of 4556 bp in length
39395 39495: gap of 100 bp
39495 43669: contig of 4175 bp in length
43670 43770: gap of 100 bp
43770 49948: contig of 6179 bp in length
49949 50049: gap of 100 bp
50049 55578: contig of 5530 bp in length
55579 61060: gap of 100 bp
61060 60959: contig of 5281 bp in length
60959

```

```

* 60960 61059: gap of 100 bp
* 61060 66394: contig of 5335 bp in length
* 66394 66494: gap of 100 bp
* 66494 72947: contig of 6453 bp in length
* 72947 73047: gap of 100 bp
* 73047 80874: contig of 7827 bp in length
* 80874 80974: gap of 100 bp
* 80974 87847: contig of 6873 bp in length
* 87847 87948: gap of 100 bp
* 87948 97279: contig of 9332 bp in length
* 97279 97379: gap of 100 bp
* 97379 106037: contig of 8658 bp in length
* 106037 106137: gap of 100 bp
* 106137 114206: contig of 8069 bp in length
* 114206 114306: gap of 100 bp
* 114306 129830: contig of 15524 bp in length
* 129830 129930: gap of 100 bp
* 129930 149800: contig of 19870 bp in length.
* 129931
FEATURES
      source
      1..149800
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="11"
      /map="11"
      /clone="RP11-11C15"
      /clone_lib="RPC1-11 Human Male BAC"
      1..1441
      /note="assembly_fragment"
      1542..2883
      /note="assembly_fragment"
      2984..4649
      /note="assembly_fragment"
      4750..6112
      /note="assembly_fragment"
      6213..7413
      /note="assembly_fragment"
      7514..9050
      /note="assembly_fragment"
      9151..10944
      /note="assembly_fragment"
      11045..13203
      /note="assembly_fragment"
      13304..16118
      /note="assembly_fragment"
      16219..18673
      /note="assembly_fragment"
      clone_end:T7
      vector_side:right"
      18774..21180
      /note="assembly_fragment"
      21281..23200
      /note="assembly_fragment"
      23301..26526
      /note="assembly_fragment"
      26627..29655
      /note="assembly_fragment"
      29756..34738
      /note="assembly_fragment"
      34839..39394
      /note="assembly_fragment"
      39495..43669
      /note="assembly_fragment"
      43770..49948
      /note="assembly_fragment"
      50049..55578
      /note="assembly_fragment"
      clone_end:SP6
      vector_side:right"
      55679..60959
      /note="assembly_fragment"
      61060..66394
      /note="assembly_fragment"

```

```
misc_feature 66495..72947
/note="assembly_fragment"
misc_feature 73048..80874
/note="assembly_fragment"
misc_feature 80975..87847
/note="assembly_fragment"
misc_feature 87948..97279
/note="assembly_fragment"
misc_feature 97380..106037
/note="assembly_fragment"
misc_feature 106138..114206
/note="assembly_fragment"
misc_feature 114307..129830
/note="assembly_fragment"
misc_feature 129931..149800
/note="assembly_fragment"
BASE COUNT 39569 a 33827 c 32852 g 40744 t 2808 others
ORIGIN
Query Match 33.9%; Score 428; DB 2; Length 149800;
Best Local Similarity 100.0%; Pred.No.1.2e-233; Indels 0; Gaps 0;
Matches 428; Conservative 0; Mismatches 0;

Qy 3 CTCCTTCGAGCTGAGCTTACCTCGGCGCAAAAGAGCGAGCGGCGGCGAGTGGAA 62
Db 19234 CTCCTTCGAGCTGAGCTTACCTCGGCGCAAAAGAGCGAGCGGCGGCGAGTGGAA 19175

Qy 63 GCTGAGATTCGCGGGTGGGCGGAGCGAGCTGCTGCTGGTCTGAGCGCGCGGAGAGC 122
Db 19174 GCTGAGATTCGCGGGTGGGCGGAGCGAGCTGCTGCTGGTCTGAGCGCGCGGAGAGC 19115

Qy 123 GGGCGGAGCGCTGATCGCTCCTCGAATCGGAGGCTCCAGTGGGGTTCCTTAGGG 182
Db 19114 GGGCGGAGCGCTGATCGCTCCTCGAATCGGAGGCTCCAGTGGGGTTCCTTAGGG 19055

Qy 183 CCCAAAGCCCGCCCGGCTCCAAAGCTCCAGGCGCTCCAGGCGACCCGCTGCTCGGC 242
Db 19054 CCCAAAGCCCGCCCGGCTCCAAAGCTCCAGGCGCTCCAGGCGACCCGCTGCTCGGC 18995

Qy 243 CCTTCCTTCGCTCAGAAAGTCGCCCTCGGCGAGTCTGTCCTCAAGGTTTCTTCGAA 302
Db 18994 CCTTCCTTCGCTCAGAAAGTCGCCCTCGGCGAGTCTGTCCTCAAGGTTTCTTCGAA 18935

Qy 303 AGAATCTGAGAGCGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 362
Db 18934 AGAATCTGAGAGCGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 18875

Qy 363 CCAGCCCGAGAGTCTGCTGCTCAATGATTTTCCCTGGCTTCTCTGCTGCTTAT 422
Db 18874 CCAGCCCGAGAGTCTGCTGCTCAATGATTTTCCCTGGCTTCTCTGCTGCTTAT 18815

Qy 423 CTACTGGG 430
Db 18814 CTACTGGG 18807

RESULT 8
AC024604 172546 bp DNA linear HTG 30-AUG-2001
LOCUS Homo sapiens chromosome Chromosome 10 clone RP11-331G19, WORKING
DEFINITION DRAFT SEQUENCE, 22 unordered pieces.
ACCESSION AC024604
VERSION AC024604.3 GI:9389428
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 172546)
REFERENCE Smith,D.R.
AUTHORS Genome Therapeutics Corporation Sequencing Center: Human Genome
TITLE Sequence Data
JOURNAL Unpublished
```

2 (bases 1 to 172546)  
Smith, D.R.  
Direct Submission  
Submitted (01-MAR-2000) Genome Therapeutics Corporation, 100 Beaver  
Street, Waltham, MA 02453, USA  
On Jun 9, 2000 this sequence version replaced gi:7549605.  
-----  
Genome Center  
Center: Genome Therapeutics Corporation  
Center code: GTC  
Web site: <http://www.genomecorp.com/>  
Contact: gtc-seqcenter@genomecorp.com  
-----  
Project Information  
Center project name: hg202  
-----  
Summary Statistics  
Sequencing Vector: N/A  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 990315  
Consensus quality: 152033 bases at least Q40  
Consensus quality: 163535 bases at least Q30  
Consensus quality: 165256 bases at least Q20  
Insert size: 170446; sum-of-contigs  
Quality coverage: 4.3x in Q20 bases; sum-of-contigs  
-----

NOTE: This is a 'working draft' sequence. It currently  
\* consists of 22 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1339: contig of 1339 bp in length  
1439: gap of unknown length  
1789: contig of 1350 bp in length  
2889: gap of unknown length  
4582: contig of 1673 bp in length  
4662: gap of unknown length  
6344: contig of 1682 bp in length  
6444: gap of unknown length  
9221: contig of 2777 bp in length  
9222: gap of unknown length  
11997: contig of 2676 bp in length  
12097: gap of unknown length  
14867: contig of 2770 bp in length  
14967: gap of unknown length  
19919: contig of 4952 bp in length  
20019: gap of unknown length  
24125: contig of 4106 bp in length  
24225: gap of unknown length  
29444: contig of 5219 bp in length  
29544: gap of unknown length  
34568: contig of 5024 bp in length  
34669: gap of unknown length  
39414: contig of 4746 bp in length  
39514: gap of unknown length  
45885: contig of 6371 bp in length  
45885: gap of unknown length  
50997: contig of 5012 bp in length  
51097: gap of unknown length  
59143: contig of 8046 bp in length  
59243: gap of unknown length  
59244: gap of unknown length  
67780: gap of unknown length  
67879: gap of unknown length  
76429: contig of 8550 bp in length  
76529: gap of unknown length  
86704: contig of 10175 bp in length  
86705: gap of unknown length  
103147: contig of 16343 bp in length  
103148: gap of unknown length  
16343: contig of 13096 bp in length  
16443: gap of unknown length  
135318: contig of 18875 bp in length  
135319: gap of unknown length



TITLE	Secreted expressed sequence tags (sESTs)
JOURNAL	Patent: JP 2001519666-A 873 23-OCT-2001;
COMMENT	GENETICS INSTITUTE INC
	PN JP 2001519666-A/873
	PD 23-OCT-2001
	PF 10-APR-1998 JP 1998543068
	PR 10-APR-1997 US 08/835913
	PI KENNETH JACOBSS, JOHN M MCCOY, EDWARD R LAVALLE, LISA A RACIE, PI DAVID MERBERG,
	PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
	C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:
	Double;
	CC Topology: Linear;
FH	Key Location/Qualifiers.
	Location/Qualifiers
	1. .471
	/organism="Zea mays"
	/mol_type="genomic DNA"
	/db_xref="taxon:4577"
BASE COUNT	93 a 126 c 143 g 108 t 1 others
ORIGIN	

	Query Match	25.7%;	Score 324;	DB 6;	Length 471;
	Best Local Similarity	100.0%;	Pred. No. 6.9e-174;		
Matches	324;	Conservative	0;	Mismatches	0; Gaps 0;
Qy	487	GCGAACCCCATGAAGCTGGGCTGCATCTCCTGATCAAGAGAGAGAGGTGGAGGCCACC	546		
Dd	139	GCGAACCCCATGAAGCTGGGCTGCATCTCCTGATCAAGAGAGAGAGGTGGAGGCCACC	198		
Qy	547	ACGGTGGTGAAATGGTTTCACAGGCCCGAGGGCGGTAAAGATTTCCTTATTTACGAGTAT	606		
Dd	199	ACGGTGGTGAAATGGTTTCACAGGCCCGAGGGCGGTAAAGATTTCCTTATTTACGAGTAT	258		
Qy	607	CAGAATGGCCACACAGAGAGGTGGAGAGCCCCCTTTACGGGGCGCTGCAGTGGGAATGGCAGC	666		
Dd	259	CAGAATGGCCACACAGAGAGGTGGAGAGCCCCCTTTACGGGGCGCTGCAGTGGGAATGGCAGC	318		
Qy	667	AAGGACTGTGAGAACGTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCTC	726		
Dd	319	AAGGACTGTGAGAACGTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCTC	378		
Qy	727	TACACCTGCATAATGTCTCCGGGAGTTTGAGTTTGAGCGGCATCGGCCCTTTGTGAAGACG	786		
Dd	379	TACACCTGCATAATGTCTCCGGGAGTTTGAGTTTGAGCGGCATCGGCCCTTTGTGAAGACG	438		
Qy	787	ACGCGGCTGATCCCCCTTAAGATG	810		
Dd	439	ACGCGGCTGATCCCCCTTAAGATG	462		

[illegible]

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan  
(E-mail: hatorigsc@gs.c.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-42-778-9923, Fax: 81-42-778-9924)

## COMMENT

----- Genome Center  
Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hatorigsc@gs.c.riken.go.jp

----- Project Information

Center project name: HumDraft11

Center clone name: RP11-158J21

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Assembly: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 15982 bases at least Q40

Consensus quality: 165758 bases at least Q30

Consensus quality: 170239 bases at least Q20

Insert size: 174431; sum-of-contigs

Quality coverage: 4.20x in Q20 bases; sum-of-contigs

-----

NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1  
13012 contig of 13012 bp in length  
13113 24677 contig of 11565 bp in length  
24778 34640 contig of 9863 bp in length  
34741 43196 contig of 8456 bp in length  
43297 50126 contig of 6830 bp in length  
50227 57725 contig of 7499 bp in length  
57826 64612 contig of 6287 bp in length  
64713 70999 contig of 5089 bp in length  
71100 76188 contig of 7292 bp in length  
76289 83580 contig of 6831 bp in length  
83681 90511 contig of 4882 bp in length  
90612 102262 contig of 6669 bp in length  
95594 105866 contig of 3504 bp in length  
102363 105866 contig of 3504 bp in length  
105967 111142 contig of 5176 bp in length  
111243 115923 contig of 4681 bp in length  
116024 121747 contig of 5724 bp in length  
121848 125532 contig of 3685 bp in length  
125828 127927 contig of 2295 bp in length  
131185 131284 contig of 1441 bp in length  
131285 132725 contig of 1441 bp in length  
132826 135566 contig of 2741 bp in length  
135667 138000 contig of 2334 bp in length  
138101 141194 contig of 3094 bp in length  
141295 143619 contig of 2325 bp in length  
143720 146448 contig of 2729 bp in length  
146549 149531 contig of 2526 bp in length  
149632 152157 contig of 2526 bp in length  
152258 155623 contig of 3366 bp in length  
155724 158435 contig of 2075 bp in length  
160610 160610 contig of 1734 bp in length  
162344 162344 contig of 100 bp in length  
164809 164809 contig of 100 bp in length  
164909 164909 contig of 100 bp in length  
167036 167036 contig of 100 bp in length  
167037 167037 contig of 100 bp in length

\* NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
13012: contig of 13012 bp in length  
13113 24677: gap of 100 bp  
13113 24677: contig of 11565 bp in length  
24778 34640: gap of 100 bp  
24778 34640: contig of 9863 bp in length  
34641 43196: gap of 100 bp  
34641 43196: contig of 8456 bp in length  
43197 50126: gap of 100 bp  
50126 57725: contig of 6830 bp in length  
50127 57725: gap of 100 bp  
50127 57725: contig of 7499 bp in length  
57826 64612: gap of 100 bp  
57826 64612: contig of 6287 bp in length  
64613 70999: gap of 100 bp  
64613 70999: contig of 5089 bp in length  
71000 76188: gap of 100 bp  
71000 76188: contig of 7292 bp in length  
76189 83580: gap of 100 bp  
76189 83580: contig of 6831 bp in length  
83581 90511: gap of 100 bp  
83581 90511: contig of 4882 bp in length  
90512 102262: gap of 100 bp  
90512 102262: contig of 6669 bp in length  
95594 105866: gap of 100 bp  
95594 105866: contig of 3504 bp in length  
102363 105866: gap of 100 bp  
102363 105866: contig of 3504 bp in length  
105967 111142: gap of 100 bp  
105967 111142: contig of 5176 bp in length  
111143 115923: gap of 100 bp  
111143 115923: contig of 4681 bp in length  
115924 121747: gap of 100 bp  
115924 121747: contig of 5724 bp in length  
121748 125532: gap of 100 bp  
121748 125532: contig of 3685 bp in length  
125828 127927: gap of 100 bp  
125828 127927: contig of 2295 bp in length  
127928 131184: gap of 100 bp  
127928 131184: contig of 1441 bp in length  
131185 131284: gap of 100 bp  
131185 131284: contig of 1441 bp in length  
131285 132725: gap of 100 bp  
131285 132725: contig of 1441 bp in length  
132826 135566: gap of 100 bp  
132826 135566: contig of 2741 bp in length  
135667 138000: gap of 100 bp  
135667 138000: contig of 2334 bp in length  
138001 141194: gap of 100 bp  
138001 141194: contig of 3094 bp in length  
141195 143619: gap of 100 bp  
141195 143619: contig of 2325 bp in length  
143620 146448: gap of 100 bp  
143620 146448: contig of 2729 bp in length  
146449 149531: gap of 100 bp  
146449 149531: contig of 2526 bp in length  
149532 152157: gap of 100 bp  
149532 152157: contig of 2526 bp in length  
152158 155623: gap of 100 bp  
152158 155623: contig of 3366 bp in length  
155624 158435: gap of 100 bp  
155624 158435: contig of 2075 bp in length  
158436 160610: gap of 100 bp  
158436 160610: contig of 1734 bp in length  
160611 162344: gap of 100 bp  
160611 162344: contig of 100 bp in length  
162345 164809: gap of 100 bp  
162345 164809: contig of 100 bp in length  
164810 164909: gap of 100 bp  
164810 164909: contig of 100 bp in length  
164910 167036: gap of 100 bp  
164910 167036: contig of 100 bp in length  
167037 167037: gap of 100 bp  
167037 167037: contig of 100 bp in length

167137 169169: contig of 2033 bp in length  
\* 169170 169269: gap of 100 bp  
\* 169270 171052: contig of 1783 bp in length  
\* 171053 171152: gap of 100 bp  
\* 171153 172333: contig of 1781 bp in length  
\* 172334 173033: gap of 100 bp  
\* 173034 174180: contig of 1147 bp in length  
\* 174181 174280: gap of 100 bp  
\* 174281 175729: contig of 1449 bp in length  
\* 175730 175829: gap of 100 bp  
\* 175830 176305: contig of 1076 bp in length  
\* 176306 177005: gap of 100 bp  
\* 177006 178431: contig of 1426 bp in length.

FEATURES

Source  
1. .178431  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q24"  
/clone="RP11-158J21"  
misc\_feature  
1. .13012  
/note="assembly\_fragment"  
misc\_feature  
13113..24677  
/note="assembly\_fragment"  
misc\_feature  
24778..34640  
/note="assembly\_fragment"  
misc\_feature  
34741..43196  
/note="assembly\_fragment"  
misc\_feature  
43297..50126  
/note="assembly\_fragment"  
misc\_feature  
50227..57725  
/note="assembly\_fragment"  
misc\_feature  
57826..64612

Query Match 20.8%; Score 262; DB 2; Length 178431;  
Best Local Similarity 99.7%; Pred. No. 3e-138;  
Matches 382; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 49 GCGCGAGTGGAGCTGAGTTCGGGGTGGCGGGAGGCGACTGCTCGGTGGCTGA 108  
Db 111412 GCGCGAGTGGAGCTGAGTTCGGGGTGGCGGGAGGCGACTGCTCGGTGGCTGA 111471  
Qy 109 GCGCGCGAGAGCGCGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGT 168  
Db 111472 GCGCGCGAGAGCGCGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGT 111531  
Qy 169 GGGGTGCTTAGGG-CCCAAGCCGCCACCGCGCTCCAAAGCTCCAGGGGCTCCCGCAG 227  
Db 111532 GGGGTGCTTAGGGGCCCAAGCCGCCACCGCGCTCCAAAGCTCCAGGGGCTCCCGCAG 111591  
Qy 228 GCACCGGTGCTGGCCCTTCCTTCGCTCAGAAAGTGCGCCCTGGGGGCGAGTTCGTCCTCA 287  
Db 111592 GCACCGGTGCTGGCCCTTCCTTCGCTCAGAAAGTGCGCCCTGGGGGCGAGTTCGTCCTCA 111651  
Qy 288 AAGGTTCTCTCAAGAGATCTGAGAGCGCGAGTCTTGACCGAGGGAATCTCTCTGTG 347  
Db 111652 AAGGTTCTCTCAAGAGATCTGAGAGCGCGAGTCTTGACCGAGGGAATCTCTCTGTG 111711  
Qy 348 TAGCCTTGGAGCGCCAGCCCGCAGAGATGCTGCTCTTCATAGATTGTTCCCGCTGGC 407  
Db 111712 TAGCCTTGGAGCGCCAGCCCGCAGAGATGCTGCTCTTCATAGATTGTTCCCGCTGGC 111771  
Qy 408 TTCTCTGCTCTTACTCTCTGGG 430  
Db 111772 TTCTCTGCTCTTACTCTCTGGG 111794

RESULT 13  
AC024604/c 172546 bp DNA linear HTG 30-AUG-2001  
LOCUS Homo sapiens chromosome Chromosome 10 clone RP11-331G19, WORKING  
DEFINITION DRAFT SEQUENCE, 22 unordered pieces.

ACCESSION AC024604  
VERSION AC024604.3 GI:8389428  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_CANCELLED.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 172546)  
AUTHORS Smith, D.R.  
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome  
Sequence Data  
JOURNAL Unpublished  
RECORD 2 (bases 1 to 172546)  
AUTHORS Smith, D.R.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-2000) Genome Therapeutics Corporation, 100 Beaver  
Street, Waltham, MA 02453, USA  
COMMENT On Jun 9, 2000 this sequence version replaced gi:7549605.

----- Genome Center  
Center: Genome Therapeutics Corporation  
Center code: GTC  
Web site: http://www.genomecorp.com/  
Contact: gtc-seqcenter@genomecorp.com  
----- Project Information  
Center project name: hg202  
----- Summary Statistics  
Sequencing vector: N/A  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 990315  
Consensus quality: 152033 bases at least Q40  
Consensus quality: 163535 bases at least Q30  
Consensus quality: 165256 bases at least Q20  
Insert size: 170446; sum-of-contigs  
Quality coverage: 4.3x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1339: contig of 1339 bp in length  
\* 1340 1439: gap of unknown length  
\* 1440 2789: contig of 1350 bp in length  
\* 2790 2889: gap of unknown length  
\* 2890 4562: contig of 1673 bp in length  
\* 4563 4682: gap of unknown length  
\* 4683 6344: contig of 1682 bp in length  
\* 6345 9222: gap of unknown length  
\* 9223 11997: contig of 2676 bp in length  
\* 11998 12097: gap of unknown length  
\* 12098 14867: contig of 2770 bp in length  
\* 14868 14967: gap of unknown length  
\* 14968 19919: contig of 4952 bp in length  
\* 19920 24126: gap of unknown length  
\* 24127 24225: gap of unknown length  
\* 24226 29444: contig of 5219 bp in length  
\* 29445 29544: gap of unknown length  
\* 29545 34568: contig of 5024 bp in length  
\* 34569 39414: contig of 4746 bp in length  
\* 39415 39514: gap of unknown length  
\* 39515 45885: contig of 6371 bp in length  
\* 45886 45986: gap of unknown length  
\* 45987 50937: contig of 5012 bp in length  
\* 50938 51098: gap of unknown length  
\* 51099 59143: contig of 8046 bp in length  
\* 59144 59243: gap of unknown length



Qy	961	TCTGAC	TACCTT	GCATCC	ATCTCG	AGAC	CAAG	AGGA	AACTCT	CGGTA	CCAG	TGGAG	AA	1020
Db	591	TCCTG	ACATC	CTTGC	CAATC	CCCAT	CTG	AGAC	CAAG	AGGA	AACTCT	CGGTA	CCAG	TGGAG
Qy	1021	TAGAAC	CAGG	AGCA	GTGTG	ACAT	GAGT	TGGC	CTG	AACAC	CTG	AGGAC	CTG	GACAT
Db	651	TAGAAC	CAGG	AGCA	GTGTG	ACAT	GAGT	TGGC	CTG	AACAC	CTG	AGGAC	CTG	GACAT
Qy	1081	TTGACG	CAATGT	CAATG	GCCAT	CAGG	AGGG							1108
Db	711	TTGACG	CAATGT	CAATG	GCCAT	CAGG	AGGG							738

REFERENCE	AUTHORS	TITLE	JOURNAL
1	...	...	...
2	...	...	...
3	...	...	...
4	...	...	...
5	...	...	...
6	...	...	...
7	...	...	...
8	...	...	...
9	...	...	...
10	...	...	...
11	...	...	...
12	...	...	...
13	...	...	...
14	...	...	...
15	...	...	...
16	...	...	...
17	...	...	...
18	...	...	...
19	...	...	...
20	...	...	...
21	...	...	...
22	...	...	...
23	...	...	...
24	...	...	...
25	...	...	...
26	...	...	...
27	...	...	...
28	...	...	...
29	...	...	...
30	...	...	...
31	...	...	...
32	...	...	...
33	...	...	...
34	...	...	...
35	...	...	...
36	...	...	...
37	...	...	...
38	...	...	...
39	...	...	...
40	...	...	...
41	...	...	...
42	...	...	...
43	...	...	...
44	...	...	...
45	...	...	...
46	...	...	...
47	...	...	...
48	...	...	...
49	...	...	...
50	...	...	...
51	...	...	...
52	...	...	...
53	...	...	...
54	...	...	...
55	...	...	...
56	...	...	...
57	...	...	...
58	...	...	...
59	...	...	...
60	...	...	...
61	...	...	...
62	...	...	...
63	...	...	...
64	...	...	...
65	...	...	...
66	...	...	...
67	...	...	...
68	...	...	...
69	...	...	...
70	...	...	...
71	...	...	...
72	...	...	...
73	...	...	...
74	...	...	...
75	...	...	...
76	...	...	...
77	...	...	...
78	...	...	...
79	...	...	...
80	...	...	...
81	...	...	...
82	...	...	...
83	...	...	...
84	...	...	...
85	...	...	...
86	...	...	...
87	...	...	...
88	...	...	...
89	...	...	...
90	...	...	...
91	...	...	...
92	...	...	...
93	...	...	...
94	...	...	...
95	...	...	...
96	...	...	...
97	...	...	...
98	...	...	...
99	...	...	...
100	...	...	...

COMMENT

Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 14833)  
Worley, K.C.

Submitted (08-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Apr 28, 2002 this sequence version replaced [gi:16117924](#).

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>

----- Project Information

Center project name: HAXY  
Center clone name: RP11-142P10

## ----- Summary Statistics

Sequencing vector: M13;

Chemistry: Dye-primer Bodipy: 14% of reads

**Chemistry:** Dye-terminator Big Dye: 86% of reads

Assembly program: phrap; version 0.990329

Consensus quality: 127260 bases at least Q40

Consensus quality: 132735 bases at least Q30

Consensus quality: 136183 bases at least Q20

Estimated insert size: 139078; sum-of-contigs equals 139078

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draftdata.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draftdata.html)).

2437:	contig of 2437 bp in length
2537:	gap of unknown length
4912:	contig of 2275 bp in length
4912:	gap of unknown length
8488:	contig of 3576 bp in length
8588:	gap of unknown length
11163:	contig of 2575 bp in length
11263:	gap of unknown length
16638:	contig of 5375 bp in length
16738:	gap of unknown length
21870:	contig of 5132 bp in length
21970:	gap of unknown length
30780:	contig of 8810 bp in length
30880:	gap of unknown length
39267:	contig of 8387 bp in length
39367:	gap of unknown length
46890:	contig of 7523 bp in length
46990:	gap of unknown length
59656:	contig of 12666 bp in length
59756:	gap of unknown length
83309:	contig of 23553 bp in length
83409:	gap of unknown length
111857:	contig of 28448 bp in length
111957:	gap of unknown length
144833:	contig of 32876 bp in length.

FEATURES source

```

2. 11055
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-142P10"

```

BASE COUNT	ORIGIN
38456 a	33920 c 34115 g 37093 t 1249 others

Query Match 17.3%; Score 218; DB 2; Length 144833;

```

Best Local Similarity 100.0%; Pred. No. 5.7e-113;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 595 ATTTACGAGTATCGGAATGGCCACAGAGAGTGGAGAGCCCTTTTCAGGGGCGCCTGCAG 654
Db |||||
Qy 8822 ATTTACGAGTATCGGAATGGCCACAGAGAGTGGAGAGCCCTTTTCAGGGGCGCCTGCAG 8881
Db |||||
Qy 655 TGGAAATGCGACGACAGGACCTGCGAGAGTGTCCATCTACTGTCTCAACGTCACTCTGAAC 714
Db |||||
Qy 8882 TGGAAATGCGACGACAGGACCTGCGAGAGTGTCCATCTACTGTCTCAACGTCACTCTGAAC 8941
Db |||||
Qy 715 GACTCTGGCCCTTACACTGCAATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCC 774
Db |||||
Qy 8942 GACTCTGGCCCTTACACTGCAATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCC 9001
Db |||||
Qy 775 TTTGTGAAGACGACGCGCTGATCCCTTAAGAGTCAC 812
Db |||||
Qy 9002 TTTGTGAAGACGACGCGCTGATCCCTTAAGAGTCAC 9039
Db |||||

```

Search completed: December 6, 2003, 00:54:02  
Job time : 4845 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 23:23:34 ; Search time 3029 Seconds  
(without alignments)  
10118.172 Million cell updates/sec

Title: US-09-936-680-4  
Perfect score: 1261  
Sequence: 1 ccccccctccagctgagc.....tgcagaactgagaagcgg 1261

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 10

Total number of hits satisfying chosen parameters: 20295047

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estio:\*  
8: em\_hc:\*  
9: gb\_estcl:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pla:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	47.6	1201	13	EX420015
2	384	30.5	1201	13	EX445002
3	348	27.6	977	13	EX452112
4	288	22.8	1069	12	BM928131

5	268	21.3	1201	9	AL534136
6	165	13.1	443	14	D44825
7	154	12.2	509	13	EX280396
8	109	8.6	972	13	BQ066875
9	70	5.6	926	13	BX411288
10	38	3.0	471	14	CB732717
11	38	3.0	472	14	CB730533
12	34	2.7	176	9	AW832826
13	33	2.6	631	14	CB577357
14	32	2.5	439	14	CB750230
15	29	2.3	636	12	BM933157
16	29	2.3	652	10	BB652801
17	29	2.3	672	10	BB614118
18	29	2.3	700	12	B1739617
19	29	2.3	705	14	CD349958
20	29	2.3	723	14	CA749311
21	29	2.3	742	14	CB526211
22	29	2.3	750	14	CD349206
23	29	2.3	780	13	BQ770528
24	29	2.3	825	13	BQ745919
25	29	2.3	845	14	CA327438
26	29	2.3	846	14	CD355879
27	29	2.3	927	13	BQ713131
28	29	2.3	950	10	BG294174
29	29	2.3	1359	11	AK076466
30	29	2.3	3549	11	AK049747
31	29	2.3	4149	11	AK049286
32	28	2.2	712	14	CB526257
33	25	2.0	436	14	CB792033
34	24	1.9	429	9	AA744587
35	24	1.9	462	28	AQ178608
36	23	1.8	189	9	AW832827
37	23	1.8	475	14	CB729080
38	23	1.8	595	13	BU368531
39	23	1.8	669	14	CB520894
40	23	1.8	683	13	BU492633
41	23	1.8	710	10	BB653495
42	23	1.8	892	13	BU368614
43	23	1.8	986	13	BU118914
44	22	1.7	518	10	BB867746
45	21	1.7	316	13	BY070150

#### ALIGNMENTS

RESULT 1

EX420015

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX420015 Homo sapiens FETAL BRAIN Homo sapiens cdna clone  
CS0DF023YA09 5-PRIME, mRNA sequence.

EX420015 GI:30646738

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 6147.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DF023YA05QPlcluster=6147.r. Contact :

Peng Liang Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600

AL534136 AL534136  
D44825 HUMSUPY279  
EX280396 EX280396  
BQ066875 AGENCOURT  
BX411288 BX411288  
CB732717 AMGNNUC:N  
CB730533 AMGNNUC:N  
AW832826 QV2-TT000  
CB577357 AMGNNUC:C  
CB750230 AMGNNUC:N  
BM933157 UI-M-BH3-  
BB652801 BB652801  
BB614118 BB614118  
B1739617 603361873  
CD349958 UI-M-FY0-  
CA749311 UI-M-FY0-  
CB526211 UI-M-FY0-  
CD349206 UI-M-FY0-  
BQ770528 UI-M-FY0-  
BQ745919 UI-M-EX0-  
CA327438 UI-M-FY0-  
CD355879 UI-M-FY0-  
BQ713131 AGENCOURT  
BG294174 602391245  
AK076466 Mus muscu  
AK049747 Mus muscu  
AK049286 Mus muscu  
CB526257 UI-M-FY0-  
CB792033 AMGNNUC:N  
AA744587 ny25906.s  
AW832827 QV2-TT000  
CB729080 AMGNNUC:N  
BU368531 603788820  
CB520894 UI-M-GH0-  
BU492633 604131658  
BB653495 BB653495  
BU368614 603790206  
BU118914 603142291  
BB867746 BB867746  
BY070150 BY070150

Paradise Avenue Genoscope sequence ID : CS0DF023AA05QPI.  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF023YA09"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/notes="Organ: brain; Vector: pCMVSPORT\_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."  
BASE COUNT 291 a 289 c 334 g 243 t 44 others  
ORIGIN  
Query Match 47.6%; Score 600; DB 13; Length 1201;  
Best Local Similarity 99.7%; Pred. No. 1e-278;  
Matches 700; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 349 AGCTTGGAGAGCGCCAGAGAGATGCTGCTCAATAGATTGTTCCCTGGCT 408  
DB 265 AGCTTGGAGAGCGCCAGAGAGATGCTGCTCAATAGATTGTTCCCTGGCT 324  
QY 409 TCTCTGCTTATCTACTGGGTGAGTGTCTGCTTCCCTGTTGTTGGAAGTGCCTG 468  
DB 325 TCTCTGCTTATCTACTGGGTGAGTGTCTGCTTCCCTGTTGTTGGAAGTGCCTG 384  
QY 469 GAGACGAGCGCTGAGGCGACCCATGAGCTGGCTGCTGCTGCTGCTGCTGCTG 528  
DB 385 GAGACGAGCGCTGAGGCGACCCATGAGCTGGCTGCTGCTGCTGCTGCTGCTG 444  
QY 529 GAGAGGTGAGGCGACCCATGAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588  
DB 445 GAGAGGTGAGGCGACCCATGAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504  
QY 589 TTCTTTATTTACAGATATCGAATGGCCACAGAGAGTGGAGGCCCTTTTCAGGGGCGC 648  
DB 505 TTCTTTATTTACAGATATCGAATGGCCACAGAGAGTGGAGGCCCTTTTCAGGGGCGC 564  
QY 649 CTGAGTGGAGTGGCAGAGAGCGCTGCTGAGAGAGTGTCTCAATCACTGTCTCAACGTCAT 708  
DB 565 CTGAGTGGAGTGGCAGAGAGCGCTGCTGAGAGAGTGTCTCAATCACTGTCTCAACGTCAT 624  
QY 709 CTGAGAGCTCTGCGCTCTACCTGCAATGTGTCTCCCGGAGTTTGAAGTTGAGGCGCAT 768  
DB 625 CTGAGAGCTCTGCGCTCTACCTGCAATGTGTCTCCCGGAGTTTGAAGTTGAGGCGCAT 684  
QY 769 CGGCGCTTTGTGAGAGAGCGCGCTGATCCCTCAAGAGTCAACGAGAGGCTGAGAG 828  
DB 685 CGGCGCTTTGTGAGAGAGCGCGCTGATCCCTCAAGAGTCAACGAGAGGCTGAGAG 744  
QY 829 GACTTCACTCTGTGGTCTCAGAAATCATGATGATCATCTTCTGTCTTCTTCAACCTG 888  
DB 745 GACTTCACTCTGTGGTCTCAGAAATCATGATGATCATCTTCTGTCTTCTTCAACCTG 804  
QY 889 TGGCTGCTCATCGAGATGATATGCTCAGAAAGTCTCAAAAGCCGAGAGGCGAGCC 948  
DB 805 TGGCTGCTCATCGAGATGATATGCTCAGAAAGTCTCAAAAGCCGAGAGGCGAGCC 864  
QY 949 CAAGAAACCGCTCTGACTACCTTGGCCATCCCTCTGAGACCAAGAGAACTCTGCGGTA 1008  
DB 865 CAAGAAACCGCTCTGACTACCTTGGCCATCCCTCTGAGACCAAGAGAACTCTGCGGTA 924  
QY 1009 CCAGTGGAGGATGAGAGGAGCGTGTGACATGAGGTGGC 1050  
DB 925 CCAGTGGAGGATGAGAGGAGCGTGTGACATGAGGTGGC 966

RESULT 2  
BX445002

LOCUS BX445002 1201 bp mRNA linear EST 15-MAY-2003  
DEFINITION BX445002 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
CS0DF023YA09 5-PRIME, mRNA sequence.  
ACCSSION BX445002.1 GI:30782286  
VERSION BX445002  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutharia; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by life technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6147.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS1AF006ZB05QPI&cluster=6147.r. Contact :  
Rong Liang Email : rliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS1AF006ZB05QPI.  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF023YA09"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/notes="Organ: brain; Vector: pCMVSPORT\_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."  
BASE COUNT 273 a 292 c 336 g 247 t 53 others  
ORIGIN  
Query Match 30.5%; Score 384; DB 13; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 2.7e-174; Indels 0; Gaps 0;  
Matches 384; Conservative 0; Mismatches 0;  
QY 349 AGCTTGGAGAGCGCCAGAGAGATGCTGCTCAATAGATTGTTCCCTGGCT 408  
DB 255 AGCTTGGAGAGCGCCAGAGAGATGCTGCTCAATAGATTGTTCCCTGGCT 314  
QY 409 TCTCTGCTTATCTACTGGGTGAGTGTCTGCTTCCCTGTTGTTGGAAGTGCCTG 468  
DB 315 TCTCTGCTTATCTACTGGGTGAGTGTCTGCTTCCCTGTTGTTGGAAGTGCCTG 374  
QY 469 GAGACGAGCGCGTGCAGAGGCAACCCATGAGCTGCTGCTCTCTGATGAGAGAG 528  
DB 375 GAGACGAGCGCGTGCAGAGGCAACCCATGAGCTGCTGCTCTCTGATGAGAGAG 434  
QY 529 GAGAGGTGAGGCGCACACGCTGGTGAATGTTCTACAGGCCGAGGCGGTGAAGAT 588  
DB 435 GAGAGGTGAGGCGCACACGCTGGTGAATGTTCTACAGGCCGAGGCGGTGAAGAT 494  
QY 589 TTCTTTATTTACAGATATCGAATGGCCACAGAGAGTGGAGGCCCTTTTCAGGGGCGC 648  
DB 495 TTCTTTATTTACAGATATCGAATGGCCACAGAGAGTGGAGGCCCTTTTCAGGGGCGC 554  
QY 649 CTGAGTGGAGTGGCAGAGAGGCTGCTGAGAGAGTGTCTGCTCAACGTCAT 708  
DB 555 CTGAGTGGAGTGGCAGAGAGGCTGCTGAGAGAGTGTCTGCTCAACGTCAT 614  
QY 709 CTGAGTGGAGTGGCAGAGAGGCTGCTGAGAGAGTGTCTGCTCAACGTCAT 614



```

595 TTCTTTATTACAGTATCGAATGGCCACACAGAGGTGGAGAGCCCC 642
|||||
AL534136 1201 bp mRNA linear EST 12-MAY-2003
LOCUS AL534136 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CSDF005Y102 5-PRIME, mRNA sequence.
ACCESSION AL534136
VERSION AL534136.2 GI:30539643
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
TITLE Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished
On Feb 13, 2001 this sequence version replaced gi:12797629.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6147.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSDF005Y102&cluster=6147.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSDF005Y102P1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDF005Y102"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 221 a 331 c 402 g 214 t 33 others
ORIGIN
Query Match 21.3%; Score 268; DB 9; Length 1201;
Best Local Similarity 99.7%; Pred. No. 3.2e-118;
Matches 388; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 134 GCCTGATCGCTCCCTGAACTGGGGAGTCCAGTGGGTCGCTTAGGGCCCAAGCCCC 193
DB 588 GCCTGATCGCTCCCTGAACTGGGGAGTCCAGTGGGTCGCTTAGGGCCCAAGCCCC 647
QY 194 CACCGCGCTCCAAAAGTCCCAAGGCGCTCCCAAGGACCGGTCGTGGGCGCTTCCTTCGG 253
DB 648 CACCGCGCTCCAAAAGTCCCAAGGCGCTCCCAAGGACCGGTCGTGGGCGCTTCCTTCGG 707
QY 254 TCAGAAAGTCGCGCCCTGGGGCAGTTGCTGCCAAGGGTTCTCCGAAAGATCTGAGA 313
DB 708 TCAGAAAGTCGCGCCCTGGGGCAGTTGCTGCCAAGGGTTCTCCGAAAGATCTGAGA 767
QY 314 GGGCGCAGTCTTGAACGAGGGAATCTCTCTGTGTAGCTTGGAAAGCGGCGAGCCCCAGA 373
DB 768 GGGCGCA-TCCTTGACCGAGGGAATCTCTCTGTGTAGCTTGGAAAGCGGCGAGCCCCAGA 826
QY 374 AGATCGCTGCTCAATAGATTGTTTCCCTCGGCTTCTCTCGTCTTATCTACTGGGTCA 433
DB 827 AGATCGCTGCTCAATAGATTGTTTCCCTCGGCTTCTCTCGTCTTATCTACTGGGTCA 886

```

```

434 GTGTCTGCTTCCCTGTGTGTGTGAACTGCCTCGAGACGAGGCGGTGCGGCAACC 493
DB 887 GTGTCTGCTTCCCTGTGTGTGTGAACTGCCTCGAGACGAGGCGGTGCGGCAACC 946
QY 494 CCATGAAGCTGCGCTGCATCTCTGCGATG 522
DB 947 CCATGAAGCTGCGCTGCATCTCTGCGATG 975
RESULT 6
D44825
LOCUS HUMSUPY279 Human brain cDNA Homo sapiens EST 20-FEB-1998
DEFINITION sequence.
ACCESSION D44825
VERSION D44825.1 GI:1572300
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 443)
TITLE Ikeda, J., Ishida, Y., Tomiyasu, H., Yamamoto, K., Bates, G.P. and
JOURNAL Transcript map of the human chromosome 4p16.3 consisting of 627
MEDLINE cDNA clones derived from 1 Mb of the Huntington's disease locus
PUBMED 97101646
8946164
COMMENT Contact: Shinji Hadano
Japan Science and Technology Corporation, NeuroGenes Project, ICORP
Univ. of Tokai School of Med.
Bohseidai, Isehara, Kanagawa 259-1193, Japan
Tel: 81-463-91-5095
Fax: 81-463-91-4993
Email: shinji@eng.med.u-tokai.ac.jp.
FEATURES
Location/Qualifiers
1..443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NF220-K"
/tissue_type="brain"
/clone_lib="Human brain cDNA"
BASE COUNT 111 a 130 c 82 g 120 t
ORIGIN
Query Match 13.1%; Score 165; DB 14; Length 443;
Best Local Similarity 99.5%; Pred. No. 1.7e-68;
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1044 AGGTGGCTGAACACCTGAGGAGCTGGACATCCCATCTTCAGCAATGTCATGCGATCAG 1103
DB 33 AGTGGCTGAACACCTGAGGAGCTGGACATCCCATCTTCAGCAATGTCATGCGATCAG 92
QY 1104 GAGGCGCCCCCAAGGGCCCCATCGCTTCCCTTCATGATCCATGTTCTGTTCAATTT 1163
DB 93 GAGGCGCCCCCAAGGGCCCCATCGCTTCCCTTCATGATCCATGTTCTGTTCAATTT 152
QY 1164 CATCCATACATCCACCTGCGCTCTGAGCTTTTCACTCTGACTTCCCTAACTCCATCAGACCT 1223
DB 153 CATCCATACATCCACCTGCGCTCTGAGCTTTTCACTCTGACTTCCCTAACTCCATCAGACCT 212
QY 1224 CTACGCCACATAGACTCTGCCAGACTGAGAGCC 1259
DB 213 CTACGCCACATAGACTCTGCCAGACTGAGAGCC 248
RESULT 7
BX280396
LOCUS BX280396 NIH_MGC_121 Homo sapiens cDNA clone linear EST 04-MAR-2003
DEFINITION

```



Full-length cDNA libraries and normalization  
Unpublished  
Genoscope - Centre National de Sequencage  
Contact: Genoscope  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6147.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSOBA10122D01\_1&cluster=6147.r.  
Contact: Feng Liang Email: fliang@lifetech.com URL:  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID: CSOBA10122D01\_CS01060\_1.  
Location/Qualifiers  
1. .926  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODF023YA09"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_libs="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."  
BASE COUNT 240 a 194 c 241 g 249 t 2 others  
ORIGIN  
Query Match 5.6%; Score 70; DB 13; Length 926;  
Best Local Similarity 100.0%; Pred. No. 1.6e-22;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1154 TTCATTTCATTCATCATACATCCACCTGCCTCTGAGCTTTCACCTCTGACCTCCCTAACTC 1213  
|||||  
Db 778 TTCATTTCATTCATCATACATCCACCTGCCTCTGAGCTTTCACCTCTGACCTCCCTAACTC 719  
|||||  
QY 1214 CATCAGACCT 1223  
|||||  
Db 718 CATCAGACCT 709  
|||||  
RESULT 10  
CB732717  
LOCUS  
DEFINITION  
CB732717  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 471)  
Angen EST Program.  
TITLE  
Unpublished  
CONTACT: Dan Fitzpatrick  
Angen, Inc  
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00118 row: d column: 10.  
Location/Qualifiers  
1. .471  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="nrp13-00118-d10"  
/tissue\_type="pituitary"  
FEATURES  
source



ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 636)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
FURNISHED  
COMMENT  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 REVERSE5.

FEATURES  
source  
1..636  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-bsq-d-04-0-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH BMAP M S4"  
/note="Vector: pT7m3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The  
NIH BMAP M S4 library is a subtracted library of a series,  
ultimately derived from a mixture of individually tagged,  
normalized libraries from ten regions of the mouse brain  
(cerebellum, brain stems, olfactory bulbs, hypothalamus,  
cortex, amygdala, basal ganglia, pineal gland, striatum,  
hippocampus) after a series of subtractions to reduce the  
representation of cDNAs from which ESTs had already been  
generated. The following serially subtracted libraries  
were generated in this process: NIH BMAP M S4,  
NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,  
NIH BMAP M S4. The subtracted library  
cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and  
NIH BMAP M S3.1 clones from which 3' ESTs had been derived  
was used as a driver in a hybridization with a pool of  
the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1  
libraries in the form of single-stranded circles. The  
remaining single-stranded circles (subtracted library)  
was purified by hydroxyapatite column chromatography,  
converted to double-stranded circles and electroporated  
into DH10B bacteria (lifetechnologies) to generate the  
NIH BMAP M S4 library. This procedure has been previously  
described (Bonaldo, Lennon and Soares, Genome Research  
6:791-806, 1996)"  
BASE COUNT 158 a 154 c 169 g 155 t  
ORIGIN

Query Match 2.3%; Score 29; DB 12; Length 636;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 610 AATGCCACAGGAGTGAGAGCCCTT 638  
DB 241 AATGCCACAGGAGTGAGAGCCCTT 269

Search completed: December 6, 2003, 01:44:39  
Job time : 3032 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 23:25:39 ; Search time 493 Seconds  
(without alignments)  
8501.133 Million cell updates/sec

Title: US-09-936-680-4  
Perfected score: 1261  
Sequence: 1 cccctccctccgagctgagc.....tgcagaaactgagaagccgg 1261

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2201672 seqs, 166179959 residues

Word size : 10

Total number of hits satisfying chosen parameters: 1800434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1261	100.0	1261	10	US-09-997-579-4 Sequence 4, Appli
2	216	17.1	4625	11	US-09-764-891-7659 Sequence 7659, Ap
3	38	3.0	645	13	US-10-029-191-21 Sequence 21, Appl
4	38	3.0	2220	10	US-09-997-579-3 Sequence 3, Appli
5	38	3.0	2632	13	US-10-029-191-22 Sequence 22, Appl
6	38	3.0	3108	13	US-10-029-191-1 Sequence 1, Appli
7	33	2.6	574	11	US-09-764-891-1293 Sequence 1293, Ap
8	24	1.9	24	12	US-10-121-755-20 Sequence 20, Appl
9	22	1.7	462	9	US-10-121-759-19 Sequence 19, Appl
10	22	1.7	621	12	US-10-027-632-15643 Sequence 15643, A
11	21	1.7	621	13	US-10-027-632-132914 Sequence 132914,
12	21	1.7	621	13	US-10-027-632-132914 Sequence 132914,
13	20	1.6	2517	10	US-09-712-363-4 Sequence 4, Appli
14	20	1.6	2884	14	US-10-219-213-3 Sequence 3, Appli
15	19	1.5	471	9	US-09-864-761-679 Sequence 679, App
16	19	1.5	625	12	US-10-027-632-86250 Sequence 86250, A

c 17	19	1.5	625	12	US-10-027-632-86251	Sequence 86251, A
c 18	19	1.5	625	13	US-10-027-632-86250	Sequence 86250, A
c 19	19	1.5	625	13	US-10-027-632-86251	Sequence 86251, A
c 20	19	1.5	637	12	US-10-027-632-217363	Sequence 217363, A
c 21	19	1.5	637	13	US-10-027-632-217363	Sequence 217363, A
c 22	19	1.5	762	12	US-10-027-632-149779	Sequence 149779, A
c 23	19	1.5	762	13	US-10-027-632-149779	Sequence 149779, A
c 24	19	1.5	782	12	US-10-027-632-29933	Sequence 29933, A
c 25	19	1.5	782	13	US-10-027-632-29933	Sequence 29933, A
c 26	19	1.5	1626	9	US-09-881-401-3	Sequence 3, Appli
c 27	19	1.5	1626	12	US-09-799-978-17	Sequence 17, Appl
c 28	19	1.5	3506	9	US-09-847-852-1	Sequence 1, Appli
c 29	19	1.5	5011	14	US-10-293-702-1	Sequence 16, Appl
c 30	19	1.5	215980	11	US-09-972-546-16	Sequence 14763, A
c 31	19	1.5	2940917	12	US-10-027-632-174763	Sequence 174763, A
c 32	19	1.5	2940917	13	US-10-027-632-174763	Sequence 5, Appli
c 33	18	1.4	247	9	US-09-912-447-5	Sequence 160, App
c 34	18	1.4	376	10	US-09-917-800A-160	Sequence 51, Appl
c 35	18	1.4	402	12	US-10-074-511-51	Sequence 271154, A
c 36	18	1.4	418	12	US-10-027-632-271154	Sequence 271155, A
c 37	18	1.4	418	12	US-10-027-632-271154	Sequence 271154, A
c 38	18	1.4	418	13	US-10-027-632-271154	Sequence 271155, A
c 39	18	1.4	418	13	US-10-027-632-271155	Sequence 271155, A
c 40	18	1.4	454	9	US-09-770-444-586	Sequence 586, App
c 41	18	1.4	496	10	US-09-917-800A-636	Sequence 636, App
c 42	18	1.4	545	12	US-10-027-632-267789	Sequence 267789, A
c 43	18	1.4	545	12	US-10-027-632-267790	Sequence 267790, A
c 44	18	1.4	545	13	US-10-027-632-267789	Sequence 267789, A
c 45	18	1.4	545	13	US-10-027-632-267790	Sequence 267790, A

## ALIGNMENTS

### RESULT 1

US-09-997-579-4  
; Sequence 4, Application US/09997579  
; Patent No. US20020113203A1  
; GENERAL INFORMATION:

; APPLICANT: Cambridge University Technical Services  
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated s  
; TITLE OF INVENTION: channel  
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses

; FILE REFERENCE: 674558-2001  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: PCT/EP00/01783  
; PRIOR APPLICATION NUMBER: 60,129,473  
; PRIOR FILING DATE: 2000-02-24

; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4

; TYPE: DNA  
; LENGTH: 1261  
; ORGANISM: Homo sapiens

; US-09-997-579-4

Query Match 100.0%; Score 1261; DB 10; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTCCCTCCGAGCTGAGCTTACCTGGCGCGCAACGAGCGAGCGCGCGCGAGTGG 60

Db 1 CCCTCCCTCCGAGCTGAGCTTACCTGGCGCGCAACGAGCGAGCGCGCGAGTGG 60

Qy 61 AAGCTGAGTTCGCGGGTGGCGGGAGCGGCACTGTCCTGCTGCTGAGCGCGCGGAGA 120

Db 61 AAGCTGAGTTCGCGGGTGGCGGGAGCGGCACTGTCCTGCTGCTGAGCGCGCGGAGA 120

Qy 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCTAGTGGGTCGCTTAG 180

Db 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCTAGTGGGTCGCTTAG 180



Db 61 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGAC 98  
|||||

## RESULT 4

US-09-997-579-3  
; Sequence 3, Application US/09997579  
; Patent No. US20020113203A1  
; GENERAL INFORMATION:  
; APPLICANT: Cambridge University Technical Services  
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodium channel  
; TITLE OF INVENTION: channel  
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses thereof  
; FILE REFERENCE: 674558-2001  
; CURRENT APPLICATION NUMBER: US/09/997,579  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: PCT/EP00/01783  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60,129,473  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 2220  
; TYPE: DNA  
; ORGANISM: rat  
US-09-997-579-3

Query Match 3.0%; Score 38; DB 10; Length 2220;  
Best Local Similarity 100.0%; Pred. No. 7.1e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGAC 473  
Db 423 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGAC 460  
|||||

## RESULT 5

US-10-029-191-22  
; Sequence 22, Application US/10029191  
; Publication No. US20020160453A1  
; GENERAL INFORMATION:  
; APPLICANT: CURTIS, ROY A.J.  
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
; FILE REFERENCE: 210147.00XX/5U1  
; CURRENT APPLICATION NUMBER: US/10/029,191  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 09/569,978  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/134,198  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 2632  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-10-029-191-22

Query Match 3.0%; Score 38; DB 13; Length 2632;  
Best Local Similarity 100.0%; Pred. No. 7e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGAC 473  
Db 138 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGAC 175  
|||||

## RESULT 6

US-10-029-191-1  
; Sequence 1, Application US/10029191  
; Publication No. US20020160453A1

; GENERAL INFORMATION:  
; APPLICANT: CURTIS, ROY A.J.  
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 210147.00XX/5U1  
; CURRENT APPLICATION NUMBER: US/10/029,191  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 09/569,978  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/134,198  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3108  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-10-029-191-1

Query Match 3.0%; Score 38; DB 13; Length 3108;  
Best Local Similarity 100.0%; Pred. No. 6.9e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGAC 473  
Db 138 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGAC 175  
|||||

## RESULT 7

US-09-764-891-1293  
; Sequence 1293, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSEN ET AL.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1293  
; LENGTH: 574  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (517)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (540)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (556)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (564)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (571)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-891-1293

Query Match 2.6%; Score 33; DB 11; Length 574;  
Best Local Similarity 100.0%; Pred. No. 3.9e-07;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1227 CGCACCATAGACTCTGCCAGAACTGAGAAGCC 1259  
Db 2 CGCACCATAGACTCTGCCAGAACTGAGAAGCC 34  
|||||

## RESULT 8

```
US-10-121-759-20/c
; Sequence 20, Application US/10121759
; Publication No. US20030194751A1
; GENERAL INFORMATION:
; APPLICANT: Dubin, Adrienne
; APPLICANT: Kaftan, Edward
; APPLICANT: Chaplan, Sandra
; APPLICANT: Brown, Sean
; TITLE OF INVENTION: Method For Identifying Modulators Of Ion Channels
; FILE REFERENCE: PRI-0007 (ORT 1534)
; CURRENT APPLICATION NUMBER: US/10/121,759
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer Oligonucleotide
US-10-121-759-20
Query Match 1.7%; Score 24; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.035; 0; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 1234 TAAGACTCTCCGAGACTGAGAAG 1257
Db 24 TAAGACTCTCCGAGACTGAGAAG 1

RESULT 9
US-10-121-759-19
; Sequence 19, Application US/10121759
; Publication No. US20030194751A1
; GENERAL INFORMATION:
; APPLICANT: Dubin, Adrienne
; APPLICANT: Kaftan, Edward
; APPLICANT: Chaplan, Sandra
; APPLICANT: Brown, Sean
; TITLE OF INVENTION: Method For Identifying Modulators Of Ion Channels
; FILE REFERENCE: PRI-0007 (ORT 1534)
; CURRENT APPLICATION NUMBER: US/10/121,759
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer Oligonucleotide
US-10-121-759-19
Query Match 1.7%; Score 22; DB 12; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.43; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 963 TGACTACCTTGGCCATCCCATCT 984
Db 1 TGACTACCTTGGCCATCCCATCT 22

RESULT 10
US-09-864-761-15643/c
; Sequence 15643, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15643
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011242.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
US-09-864-761-15643
Query Match 1.7%; Score 22; DB 9; Length 462;
Best Local Similarity 100.0%; Pred. No. 0.36; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 1149 TTCTGTTTCATTCATTCATCCAT 1170
Db 25 TTCTGTTTCATTCATTCATCCAT 4

RESULT 11
US-10-027-632-132914
; Sequence 132914, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```

PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 132914  
LENGTH: 621  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-132914

Query Match 1.7%; Score 21; DB 13; Length 621;  
Best Local Similarity 100.0%; Pred. No. 1.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0

QY 1150 TCTGTTCAATTCATTCAT 1170  
|||||  
DB 194 TCTGTTCAATTCATTCAT 214

RESULT 12  
US-10-027-632-132914  
Sequence 132914, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 132914  
LENGTH: 621  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-132914

Query Match 1.7%; Score 21; DB 13; Length 621;  
Best Local Similarity 100.0%; Pred. No. 1.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0

QY 1150 TCTGTTCAATTCATTCAT 1170  
|||||  
DB 194 TCTGTTCAATTCATTCAT 214

RESULT 13  
US-09-712-363-4/c  
Sequence 4, Application US/09712363

Patent No. US20020164588A1  
GENERAL INFORMATION:  
APPLICANT: Eisenberg, David  
APPLICANT: Rotstein, Sergio H.  
APPLICANT: Marcotte, Edward M.  
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
FILE REFERENCE: 07419-032001  
CURRENT APPLICATION NUMBER: US/09/712,363  
CURRENT FILING DATE: 2000-11-13  
PRIOR APPLICATION NUMBER: PCT/US00/02246  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,531  
PRIOR FILING DATE: 2000-02-01  
PRIOR APPLICATION NUMBER: 60/117,844  
PRIOR FILING DATE: 1999-01-29  
PRIOR APPLICATION NUMBER: 60/118,206,  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: 60/126,593  
PRIOR FILING DATE: 1999-03-26  
PRIOR APPLICATION NUMBER: 60/134,093  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 60/134,092  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 60/165,124  
PRIOR FILING DATE: 1999-11-12  
PRIOR APPLICATION NUMBER: 60/165,086  
PRIOR FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 292  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 2517  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-4

Query Match 1.8%; Score 20; DB 10; Length 2517;  
Best Local Similarity 100.0%; Pred. No. 4.1; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 500 AGCTGGCTGCATCTCTGC 519  
|||||  
DB 82 AGCTGGCTGCATCTCTGC 63

RESULT 14  
US-10-219-219-3/c  
Sequence 3, Application US/10219219  
Publication No. US20030049668A1  
GENERAL INFORMATION:  
APPLICANT: SUZUKI, Yasuhiko  
APPLICANT: ICHIHARA, Tatsuo  
APPLICANT: SATOH, No. US20030049668A1  
TITLE OF INVENTION: Method for Determining  
Quinolone Resistance of Tubercle bacilli  
FILE REFERENCE: OP1354  
CURRENT APPLICATION NUMBER: US/10/219,219  
CURRENT FILING DATE: 2002-08-15  
PRIOR APPLICATION NUMBER: JP 2001-267769  
PRIOR FILING DATE: 2001-09-04  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2884  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (266)..(2782)  
US-10-219-219-3

Query Match 1.8%; Score 20; DB 14; Length 2884;  
Best Local Similarity 100.0%; Pred. No. 4; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 500 AGCTGGCTGCTCTCTCTGC 519  
 Db 347 AGCTGGCTGCTCTCTCTGC 328

RESULT 15

US-09-864-761-679/c  
 ; Sequence 679, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 ; FILE REFERENCE: Acomica-X-1  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 679  
 ; LENGTH: 471  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; OTHER INFORMATION: MAP TO AC007076.2  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 18  
 ; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 11  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 15  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.8  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 16

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 12  
 US-09-864-761-679

Query Match 1.5%; Score 19; DB 9; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 521 TGAAGAGAGAGAGGTGGA 539  
 Db 209 TGAAGAGAGAGAGGTGGA 191

Search completed: December 6, 2003, 01:54:59  
 Job time : 497 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 17:21:16 ; Search time 20.95 Seconds  
(without alignments)  
5014.271 Million cell updates/sec

Title: US-09-936-680-4\_COPY\_1024\_1261

Perfect score: 238

Sequence: 1 aacggagcagtgatg.....tgcagaaactgagaagccgg 238

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/ECTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36.2	15.2	7218	1	US-08-232-463-14
C 2	34	14.3	526	3	Sequence 14, Appl
C 3	34	14.3	526	3	Sequence 14, Appl
C 4	34	14.3	526	4	Sequence 14, Appl
5	32.6	13.7	16389	4	Sequence 14, Appl
C 6	31.8	13.4	74	3	Sequence 19, Appl
C 7	31.8	13.4	74	4	Sequence 19, Appl
8	31.8	13.4	3693	3	Sequence 1, Appl
9	31.8	13.4	6254	3	Sequence 126, App
C 10	31.4	13.2	367	2	Sequence 9, Appl
C 11	31.2	13.1	78	3	Sequence 20, Appl
C 12	31.2	13.1	78	4	Sequence 20, Appl
C 13	30.8	12.9	248	3	Sequence 32, Appl
C 14	30.8	12.9	248	3	Sequence 32, Appl
C 15	30.8	12.9	277	3	Sequence 3, Appl
C 16	30.8	12.9	277	3	Sequence 3, Appl
C 17	30.8	12.9	3166	4	Sequence 8, Appl
C 18	30.8	12.9	152331	3	Sequence 16, Appl
19	30.8	12.9	176373	3	Sequence 17, Appl
C 20	30.4	12.8	54	4	Sequence 9, Appl
C 21	30.4	12.8	54	4	Sequence 10, Appl
C 22	30.4	12.8	70	3	Sequence 136, App
C 23	30.4	12.8	70	3	Sequence 40, Appl
C 24	30.4	12.8	70	4	Sequence 30, Appl
25	30.4	12.8	246240	2	Sequence 20, Appl
26	30.4	12.8	246240	2	Sequence 21, Appl
27	30.4	12.8	246240	2	Sequence 22, Appl

Sequence 23, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 122, Appl  
Sequence 29, Appl  
Sequence 8632, Ap  
Sequence 8631, Ap  
Sequence 157, App  
Sequence 51, Appl  
Sequence 164, App  
Sequence 58, Appl  
Sequence 165, App  
Sequence 59, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 14, Appl  
Sequence 14, Appl

28 30 12.6 11558 5 PCT-US93-06251-23  
C 29 30 12.6 319608 4 US-09-539-333D-1  
C 30 30 12.6 319608 4 US-09-679-409-1  
31 29.8 12.5 10877 4 US-09-674-311-1  
32 29.6 12.4 1001 4 US-09-641-638-122  
33 29.6 12.4 1112 4 US-07-145-002B-29  
34 29.4 12.4 1557 4 US-09-252-991A-8632  
C 35 29.4 12.4 3156 4 US-09-252-991A-8661  
C 36 29.2 12.3 76 3 US-09-358-972-157  
C 37 29.2 12.3 76 4 US-09-358-316-51  
C 38 29.2 12.3 265 3 US-09-358-972-164  
C 39 29.2 12.3 265 4 US-09-383-316-58  
C 40 29.2 12.3 269 3 US-09-358-972-165  
C 41 29.2 12.3 269 4 US-09-383-316-59  
C 42 29.2 12.3 289 3 US-09-007-005-17  
C 43 29.2 12.3 289 3 US-09-244-796-17  
C 44 29 12.2 829 3 US-08-633-993A-14  
C 45 29 12.2 829 3 US-08-844-188-14

## ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; City: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 999149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ5pt-Fls  
US-08-232-463-14

Query Match 15.2%; Score 36.2; DB 1; Length 7218;



;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack & Heuser  
;; STREET: 520 S.W. Yamhill Street, Suite 200  
;; CITY: Portland  
;; STATE: Oregon  
;; COUNTRY: U.S.A.  
;; ZIP: 97204  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/265,503B  
;; FILING DATE: March 10, 1999  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Van Rysseberghe, Pierre C.  
;; REGISTRATION NUMBER: 33,557  
;; REFERENCE/DOCKET NUMBER: OHSU 306D  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (503) 224-6655  
;; TELEFAX: (503) 295-6679  
;; TELEX: 360619  
;; INFORMATION FOR SEQ ID NO: 14:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 526 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
US-09-265-503B-14  
  
Query Match 14.3%; Score 34; DB 4; Length 526;  
Best Local Similarity 80.0%; Pred. No. 0.039;  
Matches 40; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 107 TCCCTTCATGCATCCATGTTCTGTTTCATTCATTCATCCATCCATCCACCC 156  
Db 107 TCCATTCATCCATCCATGTTCTGTTTCATCCATTCATCCATCCACCCATCCCTCC 58  
  
RESULT 5  
US-09-741-154-3  
; Sequence 3, Application US/09741154  
; Patent No. 6437110  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen M. et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; THEREOF  
; FILE REFERENCE: CLO01061  
; CURRENT APPLICATION NUMBER: US/09/741,154  
; CURRENT FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 16389  
; TYPE: DNA  
; ORGANISM: Human  
US-09-741-154-3  
  
Query Match 13.7%; Score 32.6; DB 4; Length 16389;  
Best Local Similarity 58.9%; Pred. No. 0.56;  
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
  
QY 105 CTTCCTTCATGCATCCATGTTCTGTTTCATTCATTCATCCATCCATCCCTGCTG 164  
Db 15870 CTTCATCCATCCATCCATCCATCCCTGTTTCATCCATCCACCCACCCATCCCTGTC 15929  
  
QY 165 AGCTTCACCTCTGACTCCCTACTCCATCCAGACC 199  
Db 15930 CCACCCATCCATGCATCCATCCATCCACCCACCC 15964

RESULT 6  
US-09-030-156-19/c  
; Sequence 19, Application US/09030156  
; Patent No. 6207373  
; GENERAL INFORMATION:  
; APPLICANT: Sosnowski, Ronald G.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DETERMINATION OF LENGTH  
; FILE OF INVENTION: POLYMORPHISMS IN DNA  
; FILE REFERENCE: DAVID B. MURPHY: Nanogen 221/157  
; CURRENT APPLICATION NUMBER: US/09/030,156  
; CURRENT FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 74  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-030-156-19  
  
Query Match 13.4%; Score 31.8; DB 3; Length 74;  
Best Local Similarity 76.5%; Pred. No. 0.092;  
Matches 39; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
  
QY 106 TTCCCTTCATGCATCCATGTTCTGTTTCATTCATTCATCCATCCATCCACC 156  
Db 64 TTCCCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATCC 14  
  
RESULT 7  
US-09-645-757-19/c  
; Sequence 19, Application US/09645757  
; Patent No. 6395493  
; GENERAL INFORMATION:  
; APPLICANT: Sosnowski, Ronald G.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DETERMINATION OF LENGTH  
; FILE OF INVENTION: POLYMORPHISMS IN DNA  
; FILE REFERENCE: DAVID B. MURPHY: Nanogen 256/229  
; CURRENT APPLICATION NUMBER: US/09/645,757  
; CURRENT FILING DATE: 2000-08-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 74  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-645-757-19  
  
Query Match 13.4%; Score 31.8; DB 4; Length 74;  
Best Local Similarity 76.5%; Pred. No. 0.092;  
Matches 39; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
  
QY 106 TTCCCTTCATGCATCCATGTTCTGTTTCATTCATTCATCCATCCATCCACC 156  
Db 64 TTCCCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATCC 14  
  
RESULT 8  
US-09-056-226-1  
; Sequence 1, Application US/09056226B  
; Patent No. 6177614  
; GENERAL INFORMATION:  
; APPLICANT: Colasanti, Joseph J.  
; APPLICANT: Sundaresan, Venkatesan  
; TITLE OF INVENTION: Control of Floral Induction in Plants  
; FILE REFERENCE: CSHL94-04A4  
; CURRENT APPLICATION NUMBER: US/09/056,226B  
; CURRENT FILING DATE: 1998-04-07  
; EARLIER APPLICATION NUMBER: US 09/000,640  
; EARLIER FILING DATE: 1997-12-30  
; EARLIER APPLICATION NUMBER: US 08/804,104



```
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-332-766A-9

Query Match      13.2%; Score 31.4; DB 2; Length 367;
Best Local Similarity 61.7%; Pred. No. 0.26;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 108 CCCTTCATCCATCCATTTCTGTTTCATTCATCCATCCATCCATCCCTCTCGAGC 167
DB 196 CCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 137

QY 168 TTTCACCTCTGACTCCCTAAC 188
DB 136 TTTCATCCATCCATCCATC 116

RESULT 11
US-09-030-156-20/c
; Sequence 20, Application US/09030156
; Patent No. 6207373
; GENERAL INFORMATION:
; APPLICANT: Sosnowski, Ronald G.
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DETERMINATION OF LENGTH
; FILE REFERENCE: POLYMORPHISMS IN DNA
; FILE REFERENCE: DAVID B. MURPHY: Nanogen 221/157
; CURRENT APPLICATION NUMBER: US/09/030,156
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-030-156-20

Query Match      13.1%; Score 31.2; DB 3; Length 78;
Best Local Similarity 70.0%; Pred. No. 0.15;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 97 CCCATCGCTTCCCTTCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 156
DB 73 CCTTATTTCCCTTCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 14

RESULT 12
US-09-645-757-20/c
; Sequence 20, Application US/09645757
; Patent No. 6395493
; GENERAL INFORMATION:
; APPLICANT: Sosnowski, Ronald G.
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DETERMINATION OF LENGTH
; FILE REFERENCE: POLYMORPHISMS IN DNA
; FILE REFERENCE: DAVID B. MURPHY: Nanogen 256/229
; CURRENT APPLICATION NUMBER: US/09/645,757
; CURRENT FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-645-757-20

Query Match      13.1%; Score 31.2; DB 4; Length 78;
Best Local Similarity 70.0%; Pred. No. 0.15;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 97 CCCATCGCTTCCCTTCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 156
DB 73 CCTTATTTCCCTTCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 14

RESULT 13
US-09-007-005-32/c
; Sequence 32, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 248
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-007-005-32

Query Match      12.9%; Score 30.8; DB 3; Length 248;
Best Local Similarity 24.0%; Pred. No. 0.35;
Matches 23; Conservative 43; Mismatches 30; Indels 0; Gaps 0;

QY 99 CCATCGCTTCCCTTCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 158
DB 213 YCTTGTTTTCYCYAIGTYYTGYGYYGYYTYYTYYGYYGYYGYYGYYGYYGYYGYYGYYG 154

QY 159 CCTCTGAGCTTTACCTCTGACTCCCTAATCCATC 194
DB 153 YAYCYGTYTYYTAYCYGYYGYYGYYGYYGYYGYYGYYGYYGYYGYYGYYGYYGYYGYYGYYG 118

RESULT 14
US-09-244-796-32/c
; Sequence 32, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: FUSIONS
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 248
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-244-796-32

Query Match      12.9%; Score 30.8; DB 3; Length 248;
Best Local Similarity 24.0%; Pred. No. 0.35;
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 23:19:13 ; Search time 441 Seconds  
(without alignments)  
7718.801 Million cell updates/sec

Title: US-09-936-680-4  
Perfect score: 1261  
Sequence: 1 cccctccctccgagctgagc.....tgccagaactgagagcgg 1261

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 10

Total number of hits satisfying chosen parameters: 1212755

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:	25:
/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1261	100.0	1261	21	Human beta3 cDNA.
2	1104	87.5	4052	24	Human signal trans
3	860	68.2	1510	22	Human novel sodium
4	798	63.3	978	22	Human EST-derived
5	787	62.4	1045	22	Human polynucleoti
6	396	31.4	1195	23	DNA encoding novel
7	391	31.0	3531	23	DNA encoding novel
8	324	25.7	471	20	EST clone BM4. Ho

9	218	17.3	621	23	AA86762
10	216	17.1	4625	22	AA04971
11	216	17.1	4625	23	ABL97864
C 12	91	7.2	407	22	ABA08942
C 13	91	7.2	407	22	AAK53329
14	74	5.9	509	22	AA193385
15	38	3.0	645	22	AA193385
16	38	3.0	2220	21	AA06061
17	38	3.0	2632	22	AA06062
18	38	3.0	3108	22	AA06060
19	33	2.6	574	22	AA01292
20	33	2.6	574	23	ABL96750
C 21	28	2.2	36	22	AA084148
C 22	22	1.7	462	24	AA084148
C 23	21	1.7	1990	13	AA025816
C 24	21	1.7	2352	22	AA025816
C 25	21	1.7	11122	24	AA04611
C 26	20	1.6	420	21	AA049871
C 27	20	1.6	628	22	AA049871
C 28	20	1.6	2517	22	AA049871
C 29	20	1.6	2894	25	AA049871
C 30	20	1.6	3556	22	AA049871
C 31	20	1.6	39325	22	AA049871
C 32	20	1.6	4403765	22	AA049871
C 33	20	1.6	4411529	22	AA049871
C 34	19	1.5	471	22	AA049871
C 35	19	1.5	471	22	AA049871
C 36	19	1.5	471	22	AA049871
C 37	19	1.5	471	22	AA049871
C 38	19	1.5	471	22	AA049871
C 39	19	1.5	471	22	AA049871
C 40	19	1.5	471	22	AA049871
C 41	19	1.5	471	22	AA049871
C 42	19	1.5	471	22	AA049871
C 43	19	1.5	1626	17	AA049871
C 44	19	1.5	1626	22	AA049871
C 45	19	1.5	1626	24	AA049871

#### ALIGNMENTS

#### RESULT 1

AA067837  
ID AAC67837 standard; cDNA; 1261 BP.

XX AAC67837;

DT 15-FEB-2001 (first entry)

XX Human beta3 cDNA.

Human; beta sub-unit; beta3; analgesic; anticonvulsant; cerebrotective; vasotropic; cardiac; nootropic; cytosolic; dermatological; gene therapy; voltage-gated sodium channel; pain; epilepsy; stroke; ischaemia; heart disease; Jacobsen syndrome; familial nonchromatin paraganglioma; phenylketonuria; Charcot-Marie tooth disease; ss.

OS Homo sapiens.

XX WO200063367-A1.

XX 26-OCT-2000.

XX 24-FEB-2000; 2000WO-EP01783.

XX 15-APR-1999; 99US-0129473.

PA (WARNER) WARNER LAMBERT CO.  
PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Cox F, Dixon A, Jackson A, Morgan K;  
PI

XX WPI: 2000-665241/64.  
DR P-PSDB; AAB36002.  
XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated  
PT sodium channel, and their corresponding polypeptides, useful for  
PT detecting and treating sodium channel-associated conditions, e.g. pain,  
PT epilepsy and stroke -  
XX  
XX Claim 10; Page 70-71; 88pp; English.  
XX The present sequence is given in the claims of a specification  
CC relating to a novel family of beta sub-unit proteins from a  
CC voltage-gated sodium channel. Human and rat beta sub-units, which  
CC have been collectively identified as beta3, have been isolated.  
CC The polynucleotides and polypeptides are useful for screening for  
CC agonists and antagonists of sodium channels. The agonists, antagonists,  
CC proteins and nucleic acids may be used diagnosing of treating diseases  
CC or conditions associated with voltage-gated sodium channels, e.g. pain,  
CC epilepsy, stroke, ischaemia, heart disease, Jacobsen Syndrome, Familial  
CC Nonchromaffin Paraganglioma, Phenylketonuria and Charcot Marie Tooth  
CC disease.  
XX  
XX Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 other;  
SQ

Query Match 100.0%; Score 1261; DB 21; Length 1261;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTCCCTCCGAGCTGAGCTTACCTGGGCGGCAACGAGGAGGCGGCGAGTGG 60  
DB 1 CCTCTCCCTCCGAGCTGAGCTTACCTGGGCGGCAACGAGGAGGCGGCGAGTGG 60  
QY 61 AAGCTGAGTTCGCGGCTGGGCGGAGCGAGCTGCGTGGTGGTGGTGGTGGTGG 120  
DB 61 AAGCTGAGTTCGCGGCTGGGCGGAGCGAGCTGCGTGGTGGTGGTGGTGGTGG 120  
QY 121 GCGGGCGGAGCGGCTGATCGGCTCCCTCGAAGTGGGAGGCTCCAGTGGGCTTGG 180  
DB 121 GCGGGCGGAGCGGCTGATCGGCTCCCTCGAAGTGGGAGGCTCCAGTGGGCTTGG 180  
QY 181 GGGCCAAAGCCCGCCCGGCTCCAAAGCTCCGAGGCTCCGAGGCGGCTGGTGG 240  
DB 181 GGGCCAAAGCCCGCCCGGCTCCAAAGCTCCGAGGCTCCGAGGCGGCTGGTGG 240  
QY 241 GCGCTTCTCTCGTCCGAGGAGTGGCGGCTCCGAGGCTCCGAGGCTCCGAGG 300  
DB 241 GCGCTTCTCTCGTCCGAGGAGTGGCGGCTCCGAGGCTCCGAGGCTCCGAGG 300  
QY 301 AAAGAAATCTGAGAGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGC 360  
DB 301 AAAGAAATCTGAGAGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGC 360  
QY 361 GCGCCAGCCCGAGAGTCCCTCCCTCAATAGATGTTCCCTCGGCTCTCTGTGTGTT 420  
DB 361 GCGCCAGCCCGAGAGTCCCTCCCTCAATAGATGTTCCCTCGGCTCTCTGTGTGTT 420  
QY 421 ATCTACTGGGTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
DB 421 ATCTACTGGGTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
QY 481 GTGCGAGGCAACCCCATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB 481 GTGCGAGGCAACCCCATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 541 GCGCCAGCCCGAGAGTGGTTCACAGGCGGAGGCGGCTTAAAGATTTCTTATTTTAC 600  
DB 541 GCGCCAGCCCGAGAGTGGTTCACAGGCGGAGGCGGCTTAAAGATTTCTTATTTTAC 600  
QY 601 GAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGGCGGCTTCAGTGGAAAT 660  
DB 601 GAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGGCGGCTTCAGTGGAAAT 660

QY 661 GGCAGCAGGAGCTCGAGACGCTGTCCATCACTGTGCTCAAGCTCACTCTGAAACGACTCT 720  
DB 661 GGCAGCAGGAGCTCGAGACGCTGTCCATCACTGTGCTCAAGCTCACTCTGAAACGACTCT 720  
QY 721 GGCCTCTACACCTGCAATGTGTCCCGGAGTTTGAAGTTTGAAGCGCATCGGCCCTTTGTG 780  
DB 721 GGCCTCTACACCTGCAATGTGTCCCGGAGTTTGAAGTTTGAAGCGCATCGGCCCTTTGTG 780  
QY 781 AAGACGAGCGGCTGATCCCTTAAGAGTCAACCGAGGAGGCTGGAGAGACTTCACTCT 840  
DB 781 AAGACGAGCGGCTGATCCCTTAAGAGTCAACCGAGGAGGCTGGAGAGACTTCACTCT 840  
QY 841 GTGGTCTCAGAAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
DB 841 GTGGTCTCAGAAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 GAGATGATATATTGCTACAGAAAGGTCTCAAGAGCCGAGGAGGAGGAGGAGGAGGAGG 960  
DB 901 GAGATGATATATTGCTACAGAAAGGTCTCAAGAGCCGAGGAGGAGGAGGAGGAGGAGG 960  
QY 961 TCTGACTACCTTGGCCATCCCATCTGAGAAACAGGAGAACTCTGCGGTACCACTGAGGAG 1020  
DB 961 TCTGACTACCTTGGCCATCCCATCTGAGAAACAGGAGAACTCTGCGGTACCACTGAGGAG 1020  
QY 1021 TAGAACGAGCAGTGTGACATGAGTGGCTGACACCTTGAGGAGCTGGACATCCCATG 1080  
DB 1021 TAGAACGAGCAGTGTGACATGAGTGGCTGACACCTTGAGGAGCTGGACATCCCATG 1080  
QY 1081 TTGAGCAATGTCAATGGCATCAGGAGGCGGCGCCCAAGGCGGCGCCCATCGCTTCCCTTCATGC 1140  
DB 1081 TTGAGCAATGTCAATGGCATCAGGAGGCGGCGCCCAAGGCGGCGCCCATCGCTTCCCTTCATGC 1140  
QY 1141 ATCCATTTGTTCTGTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1200  
DB 1141 ATCCATTTGTTCTGTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1200  
QY 1201 GACTCCCTTAACTCCATCAGAGCTCTACGACACATAAGACTCTGCCAGAACTGAGAGCGG 1260  
DB 1201 GACTCCCTTAACTCCATCAGAGCTCTACGACACATAAGACTCTGCCAGAACTGAGAGCGG 1260  
QY 1261 G 1261  
DB 1261 G 1261

RESULT 2  
ABA93727 ID ABA93727 standard; cDNA; 4052 BP.  
XX ABA93727;  
AC  
XX  
XX 30-APR-2002 (first entry)  
XX Human signal transduction cDNA clone amy2\_2f18.  
DE Human; foetal brain; foetal kidney; melanoma; testis; amygdala;  
XX gene therapy; ss.  
XX Homo sapiens.  
OS  
XX WO200198454-A2.  
FN  
XX 27-DEC-2001.  
PD  
XX 25-APR-2001; 2001WO-IB02050.  
PP  
XX 25-APR-2000; 2000US-199380P.  
FR  
XX (GSHU-) GERMAN HUMAN GENOME PROJECT.  
PA  
XX Wiemann S;  
PI  
XX WPI; 2002-055860/07.  
DR

DR P-PSDB; ABB05689.  
XX Human cDNA sequences and clones derived from human fetal brain, fetal  
PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic  
PT screening and therapy -  
XX Claim 1; Page 174-175; 61pp; English.  
CC The present invention describes assemblages and computer readable media  
CC comprising novel human cDNA sequences and clones derived from human  
CC fetal brain, foetal kidney, melanoma, testis and amygdala cDNA  
CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the  
CC present invention which encode the proteins given in ABB05662 to  
CC ABB05729. The human cDNA sequences and clones can be used in gene  
CC therapy. The clones may be used in a variety of applications, for  
CC example they may be used in profiling assays, for providing large arrays  
CC of human genetic material for implementing large-scale screening  
CC strategies and for treating diseases via gene therapy procedures.  
XX Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 other;  
SQ  
Query Match 87.5%; Score 1104; DB 24; Length 4052;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1254; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 CTCCTTCGAGCTGAGCTTACCTCGGCGCAACAGCGAGCGAGGCGCGAGTGGAA 62  
Db 431 CTCCTTCGAGCTGAGCTTACCTCGGCGCAACAGCGAGCGAGGCGCGAGTGGAA 490  
QY 63 GCTGAGTTCGCGGGTGGCGGGAGGAGGAGTCTCCGCTGCTGAGGCGCGGAGAGC 122  
Db 491 GCTGAGTTCGCGGGTGGCGGGAGGAGGAGTCTCCGCTGCTGAGGCGCGGAGAGC 550  
QY 123 GGGCGGAGGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTTATGGG 182  
Db 551 GGGCGGAGGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTTATGGG 610  
QY 183 CCCAAGGCCCAACCGGCTCCAAAGCTCCAGGCGCTCCCGAGGCGCTCGGCTCGGC 242  
Db 611 CCCAAGGCCCAACCGGCTCCAAAGCTCCAGGCGCTCCCGAGGCGCTCGGCTCGGC 670  
QY 243 CTTCTCTCGGTGAGAAAGTCCGCTCGGCGGAGTTCGTCCTCAAGGGTTCTCGAA 302  
Db 671 CTTCTCTCGGTGAGAAAGTCCGCTCGGCGGAGTTCGTCCTCAAGGGTTCTCGAA 730  
QY 303 AGAATCTGAGAGGCGGAGTCTTACCGAGGGAATCTCTGTGTAGCTTGAAGCG 362  
Db 731 AGAATCTGAGAGGCGGAGTCTTACCGAGGGAATCTCTGTGTAGCTTGAAGCG 790  
QY 363 CCAGGCCGAGAGTCTGCTTCAATAGATTGTTTCCCTGCTTCTCTGCTGCTTAT 422  
Db 791 CCAGGCCGAGAGTCTGCTTCAATAGATTGTTTCCCTGCTTCTCTGCTGCTTAT 850  
QY 423 CTACTGGGTGAGTCTGCTTCCCTGTGTGTGGAAGTGCCTCGAGAGCGAGGCGGT 482  
Db 851 CTACTGGGTGAGTCTGCTTCCCTGTGTGTGGAAGTGCCTCGAGAGCGAGGCGGT 910  
QY 483 GCAGGGCAACCCCATGAGCTGCGCTGATCTCTGATGAAGAGAGAGGAGTGGAGGC 542  
Db 911 GCAGGGCAACCCCATGAGCTGCGCTGATCTCTGATGAAGAGAGAGGAGTGGAGGC 970  
QY 543 CACCACGGTGGTGAATGTTCTACAGGCCCGGCGGTAAAGATTTCCTTATTACGA 602  
Db 971 CACCACGGTGGTGAATGTTCTACAGGCCCGGCGGTAAAGATTTCCTTATTACGA 1030  
QY 603 GTATCGGAATGGCCACAGGAGTGGAGAGCCCTTTTCAGGGCGGCTCGAGTGAATGG 662  
Db 1031 GTATCGGAATGGCCACAGGAGTGGAGAGCCCTTTTCAGGGCGGCTCGAGTGAATGG 1090  
QY 663 CAGCAAGACTGTCAGAGAGTGTCCATCATCTGCTCAACGTCTCAACGACTCTGG 722  
Db 1091 CAGCAAGACTGTCAGAGAGTGTCCATCATCTGCTCAACGTCTCAACGACTCTGG 1150

QY 723 CCTCTACACCTGCAATGTGTCCTGGAGTGTGAGTTTGAGCGCATCGGCTTTGTCAA 782  
Db 1151 CCTCTACACCTGCAATGTGTCCTGGAGTGTGAGTTTGAGCGCATCGGCTTTGTCAA 1210  
QY 783 GACGACGCGGTGATGTCCTTAAAGAGTACCGAGAGGCTCGAGAGGACTTCACTCTGT 842  
Db 1211 GACGACGCGGTGATGTCCTTAAAGAGTACCGAGAGGCTCGAGAGGACTTCACTCTGT 1270  
QY 843 GGTCTCAGAAATCATGATGTACATCCCTTCTGCTTCTCTCTCTCTCTCTCTCTCTCA 902  
Db 1271 GGTCTCAGAAATCATGATGTACATCCCTTCTGCTTCTCTCTCTCTCTCTCTCTCA 1330  
QY 903 GATCATATATTGCTACAGAAAGTCTCAAAAGCCGAGAGGAGGAGGAGGAGGAGGAGG 962  
Db 1331 GATCATATATTGCTACAGAAAGTCTCAAAAGCCGAGAGGAGGAGGAGGAGGAGGAGG 1390  
QY 963 TGACTACTTGGCCATCCCATCTTGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1022  
Db 1391 TGACTACTTGGCCATCCCATCTTGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1450  
QY 1023 GAACAGGAGGAGTGTGATGATGAGTGGCTTGAACACCTCTGAGGAGGAGGAGGAGGAGG 1082  
Db 1451 GAACAGGAGGAGTGTGATGATGAGTGGCTTGAACACCTCTGAGGAGGAGGAGGAGGAGG 1510  
QY 1083 CAGCAATGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTTCATGCA 1142  
Db 1511 CAGCAATGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTTCATGCA 1570  
QY 1143 CCATGTGTTGTTCAATTCATTCATATCCATATCATCCATCGCTTCTGAGCTTTCACCTCTGA 1202  
Db 1571 CCATGTGTTGTTCAATTCATTCATATCCATATCATCCATCGCTTCTGAGCTTTCACCTCTGA 1630  
QY 1203 CTCCTTAACCTCCATCAGACCTCTAAGCACCATAAGACTCTGCGAGAGGAGGAGGAGG 1259  
Db 1631 CTCCTTAACCTCCATCAGACCTCTAAGCACCATAAGACTCTGCGAGAGGAGGAGGAGG 1687  
RESULT 3  
AAF84146  
ID AAF84146 standard; cDNA; 1510 BP.  
XX AAF84146;  
AC AAF84146;  
DT 07-SEP-2001 (first entry)  
XX Human novel sodium channel beta1-like subunit encoding cDNA.  
DE Sodium channel; sensory neurone specific channel; beta1-like subunit;  
KW SNS; therapeutic; pain; analgesic; ss.  
XX Homo sapiens.  
XX  
Key Location/Qualifiers  
FT CDS  
FT 213..860  
FT /\*tag= a  
FT /product= "sodium channel beta1-like subunit"  
XX W020014293-A2.  
XX  
XX 21-JUN-2001.  
XX 14-DEC-2000; 2000WO-GB04802.  
XX 17-DEC-1999; 99GB-0029970.  
XX (GLAXO) GLAXO GROUP LTD.  
XX Plumptre M, Powell AJ, Sanseau P;  
XX WPI; 2001-398129/42.  
XX P-PSDB; AAB85206.  
XX Novel sub-unit for voltage-gated sodium channel proteins for producing

PT agents useful for treating pain -  
 PS Claim 4; Page 29-30; 31pp; English.

XX The invention provides a novel betal-like sub-unit for voltage-gated  
 CC sodium ion channel polypeptide, specifically a sensory neuron specific  
 CC channel (SNS) subunit. The novel betal-like subunit is useful for  
 CC producing a therapeutic agent which is useful treating pain in a patient.  
 CC The subunit can be expressed by standard recombinant methodology. The  
 CC present sequence represents a human novel sodium channel betal-like  
 CC subunit encoding cDNA.

XX SQ Sequence 1510 BP; 350 A; 410 C; 414 G; 336 T; 0 other;

Query Match 68.2%; Score 860; DB 22; Length 1510;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	349	AGCCTTGGAAGCGCCGAGAGATGCCCTTCAATAGATGTTTCCCTGGCT	408
Db	186	AGCCTTGGAAGCGCCGAGAGATGCCCTTCAATAGATGTTTCCCTGGCT	245
Qy	409	TCTCTGCTGTTATCTACTGGTCAGTCTCTCTTCCCTGTGTGTGAAGTGCCTCG	468
Db	246	TCTCTGCTGTTATCTACTGGTCAGTCTCTCTTCCCTGTGTGTGAAGTGCCTCG	305
Qy	469	GAGACGAGGCGGTGAGGCAACCCCATGAGCTGGCTGCAATCTCTGCAATGAGAGA	528
Db	306	GAGACGAGGCGGTGAGGCAACCCCATGAGCTGGCTGCAATCTCTGCAATGAGAGA	365
Qy	529	GAGACGAGGCGGTGAGGCAACCCCATGAGCTGGCTGCAATCTCTGCAATGAGAGA	588
Db	366	GAGACGAGGCGGTGAGGCAACCCCATGAGCTGGCTGCAATCTCTGCAATGAGAGA	425
Qy	589	TTCCTATTACGAGTATGCAATGAGGCAACGAGAGTGGAGGCGCTTTCAGGGGCGC	648
Db	426	TTCCTATTACGAGTATGCAATGAGGCAACGAGAGTGGAGGCGCTTTCAGGGGCGC	485
Qy	649	CTGAGTGGATGTCACGAGGACCTGACGAGGCTGTCATCTGCTCAACGTCAT	708
Db	486	CTGAGTGGATGTCACGAGGACCTGACGAGGCTGTCATCTGCTCAACGTCAT	545
Qy	709	CTGAGTGGATGTCACGAGGACCTGACGAGGCTGTCATCTGCTCAACGTCAT	768
Db	546	CTGAGTGGATGTCACGAGGACCTGACGAGGCTGTCATCTGCTCAACGTCAT	605
Qy	769	CGGCGCTTGTGAGAGAGCGCGCTGATCCCTTACGAGTCCAGGAGGCTGGAG	828
Db	606	CGGCGCTTGTGAGAGAGCGCGCTGATCCCTTACGAGTCCAGGAGGCTGGAG	665
Qy	829	GACTTCACCTCTGTGGTCTCAGAAATCATGATGATACCTCTTCTGCTTCTCACCCTG	888
Db	666	GACTTCACCTCTGTGGTCTCAGAAATCATGATGATACCTCTTCTGCTTCTCACCCTG	725
Qy	889	TGGTCTGCTCATCGAGATGATATTTGTACAGAAAGTCTCAAAGCCGAGAGGAGCC	948
Db	726	TGGTCTGCTCATCGAGATGATATTTGTACAGAAAGTCTCAAAGCCGAGAGGAGCC	785
Qy	949	CAAGAAAACGCTCTGACTACCTTGCATCCATCTGAGAAACAGAGGAACTCTGGGTA	1008
Db	786	CAAGAAAACGCTCTGACTACCTTGCATCCATCTGAGAAACAGAGGAACTCTGGGTA	845
Qy	1009	CCAGTGGAGGAAATGACACAGAGCAGTGTGACATGAGGTGGCTGACACCTGAGGACT	1068
Db	846	CCAGTGGAGGAAATGACACAGAGCAGTGTGACATGAGGTGGCTGACACCTGAGGACT	905
Qy	1069	GGACATCCCATGTTTTCAGCAATGTCATATGGCATCAGGAGGCGCGCCCAAGGGCCCATCGC	1128
Db	906	GGACATCCCATGTTTTCAGCAATGTCATATGGCATCAGGAGGCGCGCGCCCAAGGGCCCATCGC	965
Qy	1129	TTCCTTTCATGATCCATGTTTTCAGCAATGTCATATGTCATTCATCATCATCATCATCAT	1188
Db	966	TTCCTTTCATGATCCATGTTTTCAGCAATGTCATATGTCATTCATCATCATCATCATCAT	1025

Qy 1189 GCCTTCACCTCTGACTCCCTAACTCCATCAGACTCTTACGACCAATAGACTCTGCAGA 1248  
 Db 1026 GCCTTCACCTCTGACTCCCTAACTCCATCAGACTCTTACGACCAATAGACTCTGCAGA 1085

Qy 1249 ACTGAGAAGCC 1259  
 Db 1086 ACTGAGAAGCC 1096

#### RESULT 4

AAH98320  
 ID AAH98320 standard; cDNA; 978 BP.

AC AAH98320;

XX 12-OCT-2001 (first entry)

XX Human EST-derived coding sequence SEQ ID NO: 177.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KW diagnosis; forensic test; gene mapping; genetic disorder;  
 KW biodiversity; gene therapy; nutrition; ss.

OS Homo sapiens.

XX WO200154477-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02687.

XX 25-JAN-2000; 2000US-0491404.

XX 17-JUL-2000; 2000US-0617746.

XX 03-AUG-2000; 2000US-0631451.

XX 15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;

PI Cao Y, Drmanac RA, Zhang J, Werhman T;

XX MPI; 2001-476164/51.

XX P-PSDB; AAM23661.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising

PT antibodies and research use -

XX Claim 1; Page 299-300; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC from the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
 CC of the invention.

XX SQ Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 other;

Query Match 63.3%; Score 798; DB 22; Length 978;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 GCGGACTCTCGCTGCTGCTGAGCGCGCGGAGCGCGCGGAGCGCGCTGATCGGCTCC 147  
 Db 7 GCGGACTCTCGCTGCTGCTGAGCGCGCGGAGCGCGCGGAGCGCGCTGATCGGCTCC 66

Qy 148 CTCGAGCTGGGGAGGTCAGTGGGCTCGCTTAGGGCCCCAAAGCCCCCGGCTCCAAA 207  
 Db 67 CTCGAGCTGGGGAGGTCAGTGGGCTCGCTTAGGGCCCCAAAGCCCCCGGCTCCAAA 126

QY	208	AGCTCCAGGGGCTCTCCACAGGACCGGTCTCGGCCCTTCTTGGGTGAGAAAGTCGGCC	267
DB	127	AGCTCCAGGGGCTCTCCAGGACCGGTCTCGGCCCTTCTTGGGTGAGAAAGTCGGCC	186
QY	268	CTTGGGGGAGTTGTGTCCTGCAAGAGGTTCTCGAAAGAAATCTGAGAGGCGCAGTCTCTTG	327
DB	187	CTTGGGGGAGTTGTGTCCTGCAAGAGGTTCTCGAAAGAAATCTGAGAGGCGCAGTCTCTTG	246
QY	328	ACCGAGGGAATCTCTCTGTGTAGCTTGTGAAGCCGCGACGCCCCAGAGATGCTTGCCTTC	387
DB	247	ACCGAGGGAATCTCTCTGTGTAGCTTGTGAAGCCGCGACGCCCCAGAGATGCTTGCCTTC	306
QY	388	AATAGATTGTTTCCCTCGCTTCTCTCGTGCTTATCTACTGGGTCAGTCTGTCTTCCCT	447
DB	307	AATAGATTGTTTCCCTCGCTTCTCTCGTGCTTATCTACTGGGTCAGTCTGTCTTCCCT	366
QY	448	GTGCTGTGGAGTGCCCTCGAGAGCGGAGGCGGTGCAGGGCAACCCCATGAAGCTGGGC	507
DB	367	GTGCTGTGGAGTGCCCTCGAGAGCGGAGGCGGTGCAGGGCAACCCCATGAAGCTGGGC	426
QY	508	TGCAPCTCCTGCATGAAGAGAGAGGAGGTGGAGGCCACACCGTGTGGAAATGGTTCATC	567
DB	427	TGCATCTCTGCATGAAGAGAGAGGAGGTGGAGGCCACACCGTGTGGAAATGGTTCATC	486
QY	568	AGGCCCGAGGGCGGTAAAGATTTCCTTATTTACGAGTATCGAAATGGCCACAGAGGTG	627
DB	487	AGGCCCGAGGGCGGTAAAGATTTCCTTATTTACGAGTATCGAAATGGCCACAGAGGTG	546
QY	628	GAGAGCCCTTTTCAGGGGCGCTGTCAGTGGAAATGCGCAGCAGGACCTGCAGAGCTGTCC	687
DB	547	GAGAGCCCTTTTCAGGGGCGCTGTCAGTGGAAATGCGCAGCAGGACCTGCAGAGCTGTCC	606
QY	688	ATCAGTGTGCTCAGCTCAGTGTGAGAGGAGTTCCTTATTTACGAGTATCGAAATGGCCACAGAGGTG	747
DB	607	ATCAGTGTGCTCAGCTCAGTGTGAGAGGAGTTCCTTATTTACGAGTATCGAAATGGCCACAGAGGTG	666
QY	748	GAGTTTGATTGAGGGCGCATCGGCCCTTTGTGTGAGAGCAGCGCGCTCATCCCCCTAAGA	807
DB	667	GAGTTTGATTGAGGGCGCATCGGCCCTTTGTGTGAGAGCAGCGCGCTCATCCCCCTAAGA	726
QY	808	GTCACCGAGGAGGCTGGAGAGGACTTCACCTCTGTGTTCTGAGAAATCATGATGTACATC	867
DB	727	GTCACCGAGGAGGCTGGAGAGGACTTCACCTCTGTGTTCTGAGAAATCATGATGTACATC	786
QY	868	CTTCTGTGTCCTTCTCACC	885
DB	787	CTTCTGTGTCCTTCTCACC	804

## RESULT 5

AAK52345

ID AAK52345 standard; cDNA; 1045 BP

XX  
AC  
AAK52345.

**SECRET**

DT 06-NOV-2001 (first entry)

XX  
DE  
Human polynucleotide seq in NO 89XX  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  

Human; cytokine; cell proliferation

KW vaccine; peptide therapy; stem cell tissue growth factors; immunomodulation

KW nervous system disorder: arthritis  
KW growth factor; immunomodulation

XX

OS  
xx  
Homo sapiens.

PN WO200157190-A2

XX  
XX

PD 09-AUG-2001.  
vv

05-FEB-2001: 2001WO-IIS04098

[illegible]

PR	03-FEB-2000; 2000US-0436914.	
PR	27-APR-2000; 2000US-0560875.	
PR	20-JUN-2000; 2000US-0598075.	
PR	19-JUL-2000; 2000US-0620325.	
PR	01-SEP-2000; 2000US-0654936.	
PR	15-SEP-2000; 2000US-0663561.	
PR	20-OCT-2000; 2000US-0693561.	
PR	30-NOV-2000; 2000US-0728422.	
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;	
PI	Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;	
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	
XX		
DR	WPI; 2001-476283/51.	
DR	P-PSDB; AAM79212.	
XX		
PT	Nucleic acids encoding polypeptides with cytokine-like activities,	
PT	useful in diagnosis and gene therapy -	
XX		
PS	Claim 1; Page 2934-2935; 6221pp; English.	
XX		
CC	The invention relates to polynucleotides (AAKS1456-AAKS34435) and the	
CC	encoded polypeptides (AAM78333-AAM80302) that exhibit activity elating to	
CC	cytokine, cell proliferation or cell differentiation or which may induce	
CC	production of other cytokines in other cell populations. The	
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or	
CC	peptide therapy. The polypeptides have various cytokine-like activities,	
CC	e.g. stem cell growth factor activity, haematopoiesis regulating	
CC	activity, tissue growth factor activity, immunomodulatory activity and	
CC	activin/inhibin activity and may be useful in the diagnosis and/or	
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and	
CC	inflammation.	
CC	Note: Records for SEQ ID NO 2110 (AAKS2581), 2111 (AAKS2582) and 3666	
CC	(AAM80020) are omitted as the relevant pages from the sequence listing	
CC	were missing at the time of publication.	
XX		
SQ	Sequence 1045 BP; 222 A; 289 C; 314 G; 220 T; 0 other;	

Query Match	62.4%; Score 787; DB 22; Length 1045;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 837; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	124 GGCGCGAGCGGCTGATCGGCTCCCTCGAACCTGGGAGGTCAGTGGGGTCGCTTAGGCG 183
Db	50 GGCGCGGAGCGGCTGATCGGCTCCCTCGAACCTGGGAGGTCAGTGGGGTCGCTTAGGCG 109
Qy	184 CCAAGCGCCGCCACCCGGCTCCAAAAGCTCCGAGGGCTCCCGAGGACCGGTCGCTGGCC 243
Db	110 CCAAGCGCCGCCACCCGGCTCCAAAAGCTCCAGGGCTCCCGAGGACCGGTCGCTGGCC 169
Qy	244 CTTCCCTTCGGTCAGAAAGTCGCCCTCGGGCGAGTTCGTCCTCAAGGGTTCTCCGAAA 303
Db	170 CTTCCCTTCGGTCAGAAAGTCGCCCTCGGGCGAGTTCGTCCTCAAGGGTTCTCCGAAA 229
Qy	304 GAATCTGAGGCGCGCAGTCTTGACCGAGGAATCTCTCTGTGTAGGCTTTGGAGCCGC 363
Db	230 GAATCTGAGGCGCGCAGTCTTGACCGAGGAATCTCTCTGTGTAGGCTTTGGAGCCGC 289
Qy	364 CAGCCCCAGAAATGCGTCCCTCAATAGATGTTTCCCGTGGCTTCTCTGTGCTTATC 423
Db	290 CAGCCCCAGAAATGCGTCCCTCAATAGATGTTTCCCGTGGCTTCTCTGTGCTTATC 349
Qy	424 TACTGGGTCAAGTCTGCTTCCCTGTGTGTGTGAAAGTGCCTCCGAGACGAGAGGCCGTG 483
Db	350 TACTGGGTCAAGTCTGCTTCCCTGTGTGTGTGAAAGTGCCTCCGAGACGAGAGGCCGTG 409
Qy	484 CAGGCGCAACCCCATCAAGCTGGGTGCATCTCCATGAAGAGAGAGAGAGGTGGAGGCC 543
Db	410 CAGGCGCAACCCCATCAAGCTGGGTGCATCTCCATGAAGAGAGAGAGAGGTGGAGGCC 469
Qy	544 ACCAGGTGGTGAATGTTCTACAGGCCCGGAGGGCGGTAAAGATTTCTTATTTAGGAG 603

[illegible]

RESULT 6  
AAS86764  
ID AAS86764 standard; cDNA; 1195 BP.  
XX  
XX AAS86764;  
AC  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX DNA encoding novel human diagnostic protein #22568.  
DE  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
KW

New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity -  
 Claim 1; SEQ ID No 22568; 103pp; English.  
 The invention relates to isolated polynucleotide (I) and  
 polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 and gene mapping, and in recombinant production of (II). The  
 polynucleotides are also used in diagnostics as expressed sequence tags  
 for identifying expressed genes. (I) is useful in gene therapy techniques  
 to restore normal activity of (II) or to treat disease states involving

(II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AA564157-AA594564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct/sequences](http://ftp.wipo.int/pub/published/pct/sequences).

SQ	Sequence	1195 BP; 253 A; 356 C; 381 G; 205 T; 0 other;
Query Match	31.4%; Score 396; DB 23; Length 1195;	
Best Local Similarity	100.0%; Pred. No. 2.5e-182;	
Matches	396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	490 AACCCCATGAGTGGCTGCATCTCCTGCATGAAGAGAGAGAGGTGGAGGCCACCAAG 549	
Db	AAACCCCATGAGTGGCTGCATCTCCTCATGAAGAGAGAGAGGTGGAGGCCACCAAG 640	
QY	550 GTGGTGGAATGGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTTAATTTACGAGTATCGG 609	
Db	GTTGGTGGAATGGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTTAATTTACGAGTATCGG 700	
QY	610 AATGGCCACAGAGGTGAGAGAGCCCTTTTCAGGGCGGCGCTGCAGTGGAAATGGCAACAAG 669	
Db	AATGGCCACAGAGGTGAGAGAGCCCTTTTCAGGGCGGCGCTGCAGTGGAAATGGCAACAAG 760	
QY	670 GACCTCGACGAGGTGTCATCACTGTGCTCAAGTCACTCTGACACGACTCTGCGCTCTAC 729	
Db	GACCTCGACGAGGTGTCATCACTGTGCTCAAGTCACTCTGACACGACTCTGCGCTCTAC 820	
QY	730 ACCTGCAATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGACGACG 789	
Db	ACCTGCAATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGACGACG 880	
QY	790 CGSCTCATCCCCCTAAGAGTCAACCGAGGAGGCTGGAGAGGACTTCACTCTGTGGTCTCA 849	
Db	CGSCTCATCCCCCTAAGATCACCGAGGAGGCTGGAGAGGACTTCACTCTGTGGTCTCA 940	
QY	850 GAATCATCATGATPACATCCTTCTGTGCTTCCTCAACC 885	
Db	GAATCATCATGATPACATCCTTCTGTGCTTCCTCAACC 976	

RESULT 7	
AAS86763	
ID	AAS86763 standard; cDNA; 3531 BP.
XX	
XX	AAS86763;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #22567.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
OS	Homo sapiens.



QY 727 TACACTGCAATGTGTCCCGGAGTTGAGTTGAGGCGCATCGCCCTTTGTGAAGACG 786  
Db 379 TACACTGCAATGTGTCCCGGAGTTGAGTTGAGGCGCATCGCCCTTTGTGAAGACG 438  
QY 787 ACGCGGCTGATCCCTTAAGAGTC 810  
Db 439 ACGCGGCTGATCCCTTAAGAGTC 462

## RESULT 9

AAS86762  
ID AAS86762 standard; cDNA; 621 BP.

AC AAS86762;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #22566.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG22575.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

XX Claim 1; SEQ ID No 22566; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 621 BP; 156 A; 155 C; 175 G; 135 T; 0 other;

XX Query Match 17.3%; Score 218; DB 23; Length 621;

Best Local Similarity 100.0%; Pred. No. 1e-95;  
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 595 ATTTACGAGTATCGAATGGCCACACGAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCAG 654

Db 58 ATTTACGAGTATCGAATGGCCACACGAGAGGTGGAGAGCCCTTTTCAGGGCGCTGCAG 117

QY 655 TGGAAATGGCAGCAAGGACCTCGAGGACGTGCCATCACTGTCTCAACGTCACCTCTGAAC 714

Db 118 TGGAAATGGCAGCAAGGACCTCGAGGACGTGCCATCACTGTCTCAACGTCACCTCTGAAC 177

QY 715 GACTCTGGCCCTCTACACTCAATGTGTCCCGGAGTTGAGTTGAGGCGCATCGGCC 774

Db 178 GACTCTGGCCCTCTACACTCAATGTGTCCCGGAGTTGAGTTGAGGCGCATCGGCC 237

QY 775 TTTGTGAAGACGACGCGGCTGATCCCTTAAGAGTCAC 812

Db 238 TTTGTGAAGACGACGCGGCTGATCCCTTAAGAGTCAC 275

## RESULT 10

AAL04971

ID AAL04971 standard; DNA; 4625 BP.

AC AAL04971;

XX 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 7659.

XX Human; reproductive system related antigen; reproductive system disorder;  
XX cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01339.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 19-APR-2000; 2000US-0198123.

XX 07-JUN-2000; 2000US-0205515.

XX 28-JUN-2000; 2000US-0209467.

XX 30-JUN-2000; 2000US-0214886.

XX 07-JUL-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 11-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0217496.

XX 26-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 14-AUG-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 22-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0223343.  
PR 01-SEP-2000; 2000US-0223344.  
PR 01-SEP-2000; 2000US-0223345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234998.  
PR 27-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465570/50.

Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition -

Disclosure; SEQ ID NO 7659; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a genomic sequence encoding a protein of the invention.

Sequence 4625 BP; 1222 A; 1130 C; 906 G; 1367 T; 0 other;

Query Match 17.1%; Score 216; DB 22; Length 4625;  
Best Local Similarity 100.0%; Pred. No. 9.1e-95;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1044 AGGTGGCCTGAACACCTGAGGAGCTGACATCCCATGTTTCAGCAATGTCATGGCATCAG 1103

Db 33 AGGTGGCCTGAACACCTGAGGAGCTGACATCCCATGTTTCAGCAATGTCATGGCATCAG 92

QY 1104 GAGGGGCCCCAAGGGCCCCATCGTTCCTTCATGATCCATTCATGTTCTGTTTCATTCATT 1163

Db 93 GAGGGGCCCCAAGGGCCCCATCGTTCCTTCATGATCCATTCATGTTCTGTTTCATTCATT 152

QY 1164 CATCCATACATCCATCGCTCTGAGCTTTTCACCTTCCTGACTCCCTTAACCTCCATCAGACCT 1223

Db 153 CATCCATACATCCATCGCTCTGAGCTTTTCACCTTCCTGACTCCCTTAACCTCCATCAGACCT 212

QY 1224 CTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCC 1259

Db 213 CTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCC 248

RESULT 11

ABL97864  
ID ABL97864 standard; DNA; 4625 BP.  
XX AC ABL97864;  
XX DT 21-JUN-2002 (first entry)  
XX DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2516.  
XX KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ds.  
XX OS Homo sapiens.  
XX PN WO20015317-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01329.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184654.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229809.  
PR 05-SEP-2000; 2000US-0229813.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231143.  
PR 08-SEP-2000; 2000US-0231144.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 23-SEP-2000; 2000US-0234997.  
PR 23-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-483232/52.  
XX Nucleic acids encoding 973 human testicular antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating testicular cancer -  
XX Disclosure; SEQ ID NO 2516; 766pp; English.  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention.  
XX Sequence 4625 BP; 1222 A; 1130 C; 906 G; 1367 T; 0 other;  
SQ  
Query Match 17.1%; Score 216; DB 23; Length 4625;  
Best Local Similarity 100.0%; Pred. No. 9.1e-95; Indels 0; Gaps 0;  
Matches 216; Conservative 0; Mismatches 0;  
QY 1044 AGGTGGCCCTGACACCTGAGGACCTGGACATCCCATGTTTCAGCATGTCATGTCATCAG 1103  
DB 33 AGGTGGCCCTGACACCTGAGGACCTGGACATCCCATGTTTCAGCATGTCATGTCATCAG 92  
QY 1104 GAGGGGCGCCGAGGCGCCGCTCCCTTCATGATCCATGTTCTGTCATTCATT 1163  
DB 93 GAGGGGCGCCGAGGCGCCGCTCCCTTCATGATCCATGTTCTGTCATTCATT 152  
QY 1164 CATCCATATACATCACCCTGCTGAGCTTTTCACCTTCGACTCCCTTAACCTCCATCAGACCT 1223  
DB 153 CATCCATATACATCACCCTGCTGAGCTTTTCACCTTCGACTCCCTTAACCTCCATCAGACCT 212  
QY 1224 CTACGACCATGATGATCTGCGCAGAACTGAGAGCC 1259  
DB 213 CTACGACCATGATGATCTGCGCAGAACTGAGAGCC 248  
RESULT 12  
ABA08942/C  
ID ABA08942 standard; cDNA; 407 BP.  
XX ABA08942;  
XX  
XX 11-JAN-2002 (first entry)  
XX Human voltage-gated Na channel subunit homologue cDNA, SEQ ID NO:718.  
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;  
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
XX chronic inflammatory condition; proliferative retinopathy;  
XX atherosclerosis; coronary heart disease; arterial ischaemia;  
XX bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; anti-inflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antidiabetic; antitubercular;  
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
KW antifungal; vulnary; antiulcer; ss.  
OS Homo sapiens.  
XX WO200157188-A2.  
XX 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US03800.  
XX 03-FEB-2000; 2000US-0496914.  
XX 27-APR-2000; 2000US-0560875.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
PI WPI; 2001-457740/49.  
XX P-PSDB; ABB11698.  
XX Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
XX e.g. arthritis and cancer -  
XX Claim 1; Page 656; 1963pp; English.  
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, and hence  
CC giving an insight into their probable biological activities, and thereby  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation and metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.  
XX SQ Sequence 407 BP; 120 A; 91 C; 86 G; 110 T; 0 other;  
Query Match 7.2%; Score 91; DB 22; Length 407;  
Best Local Similarity 99.3%; Pred. No. 6.4e-34;  
Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC were missing at the time of publication.

XX Sequence 407 BP; 120 A; 91 C; 86 G; 110 T; 0 other;

SQ Query Match 7.2%; Score 91; DB 22; Length 407;

Best Local Similarity 99.3%; Pred. No. 6.4e-34;

Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 820 GCTGGAGAGACTTCACTCTGCTGCTCAGAAATCATGATGATCATCTCTGCTCTTC 879

DB 230 GCTGGAGAGAGACTTCACTCTGCTGCTCAGAAATCATGATGATCATCTCTGCTCTTC 171

QY 880 CTCACCTTGTGGCTGCTCATCGAGATGATATTTGCTACAGAAAGGTCTCAAAAGCCGAA 939

DB 170 CTCACCTTGTGGCTGCTCATCGAGATGATATTTGCTACAGAAAGGTCTCAAAAGCCGAA 111

QY 940 GAGCGAGCCCAAGAAACCGGT 961

DB 110 GAGCGAGCCCAAGAAACCGGT 89

RESULT 14

AAI93385

ID AAI93385 standard; cDNA; 509 BP.

XX AC AAI93385;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 13445.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorders; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PR 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX DR WPI; 2001-476283/51.

XX DR P-PSDB; AAM80196.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,

XX PT useful in diagnosis and gene therapy -

XX PS Claim 1; Page 5033; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation.

XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

XX CC (AAM80020) are omitted as the relevant pages from the sequence listing

CC were missing at the time of publication.

XX Sequence 407 BP; 120 A; 91 C; 86 G; 110 T; 0 other;

SQ Query Match 7.2%; Score 91; DB 22; Length 407;

Best Local Similarity 99.3%; Pred. No. 6.4e-34;

Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 820 GCTGGAGAGACTTCACTCTGCTGCTCAGAAATCATGATGATCATCTCTGCTCTTC 879

DB 230 GCTGGAGAGAGACTTCACTCTGCTGCTCAGAAATCATGATGATCATCTCTGCTCTTC 171

QY 880 CTCACCTTGTGGCTGCTCATCGAGATGATATTTGCTACAGAAAGGTCTCAAAAGCCGAA 939

DB 170 CTCACCTTGTGGCTGCTCATCGAGATGATATTTGCTACAGAAAGGTCTCAAAAGCCGAA 111

QY 940 GAGCGAGCCCAAGAAACCGGT 961

DB 110 GAGCGAGCCCAAGAAACCGGT 89

RESULT 14

AAI93385

ID AAI93385 standard; cDNA; 509 BP.

XX AC AAI93385;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 13445.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorders; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PR 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX DR WPI; 2001-476283/51.

XX DR P-PSDB; AAM80196.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,

XX PT useful in diagnosis and gene therapy -

XX PS Claim 1; Page 5033; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation.

XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

XX CC (AAM80020) are omitted as the relevant pages from the sequence listing

Job time : 447 secs

```
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 509 BP; 88 A; 174 C; 176 G; 71 T; 0 other;

Query Match      5.9%; Score 74; DB 22; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.2e-25;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 AGCCTTGGAAGCCGCCAGCCAGAGATGCTTCAATAGATGTTTCCCTGGCT 408
   |||||
Db 436 AGCCTTGAAGCCGCCAGCCAGAGATGCTTCAATAGATGTTTCCCTGGCT 495
   |||||

QY 409 TCTCTCGTCTTAT 422
   |||||
Db 496 TCTCTCGTCTTAT 509
   |||||

RESULT 15
AAC90601
ID AAC90601 standard; cDNA; 645 BP.
XX
AC AAC90601;
XX
DT 13-MAR-2001 (first entry)
XX
DE Rat sodium channel beta3 protein Alrxa94h5 partial coding sequence.
XX
KW Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;
KW neurodegenerative disorder; mood disorder; muscle contraction; ss.
XX
OS Rattus sp.
XX
PN W0200069912-A1.
XX
PD 23-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US13144.
XX
PR 14-MAY-1999; 99US-0134198.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ;
XX
PS WPI; 2001-122743/13.
XX
PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
PT root ganglion cDNA library for use in chromosome mapping, forensic
PT medicine, monitoring clinical trials and therapeutics -
XX
PS Claim 1; Page 40-41; 145pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC rat sodium channel beta3 protein, designated Alrxa94h5. This protein is
CC involved in the generation of pain and other sensory or perceptible nerve
CC impulses, in the establishment and endurance of mood, neurodegenerative
CC and sleep disorders, and in the control of muscle contraction, including
CC movements such as the heartbeat, digestion and vascular tone. The
CC sequences can be used in predictive medicine, screening and diagnostic
CC assays, and in pharmacogenomics.
XX
SQ Sequence 645 BP; 155 A; 154 C; 181 G; 155 T; 0 other;

Query Match      3.0%; Score 38; DB 22; Length 645;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GTCTGTTCCCTGTGTGTGAGTGGCCCTCGGAGAC 473
   |||||
Db 61 GTCTGTTCCCTGTGTGTGAGTGGCCCTCGGAGAC 98
   |||||
```

Search completed: December 5, 2003, 23:33:12

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 5, 2003, 23:24:09 ; Search time 105 Seconds  
(without alignments)  
5300.801 Million cell updates/sec

Title: US-09-936-680-4  
Perfect score: 1261  
Sequence: 1 cccctccctccagctgagc.....tgccagaaactgagaagccgg 1261

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 10

Total number of hits satisfying chosen parameters: 189644

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.6	4403765	3 US-09-103-840A-2	Sequence 2, Appli
2	20	1.6	4411529	3 US-09-103-840A-1	Sequence 1, Appli
3	19	1.5	1626	1 US-08-381-433A-3	Sequence 3, Appli
4	18	1.4	344	3 US-09-157-177-128	Sequence 128, App
5	18	1.4	466	4 US-09-280-116-14	Sequence 14, Appl
6	18	1.4	713	4 US-09-535-008-44	Sequence 44, Appl
7	18	1.4	933	4 US-09-252-991A-9940	Sequence 9940, Ap
8	18	1.4	1023	4 US-09-252-991A-10137	Sequence 10137, A
9	18	1.4	3866	4 US-09-023-942A-27	Sequence 27, Appl
10	18	1.4	4506	4 US-08-178-257-1	Sequence 1, Appli
11	18	1.4	44453	4 US-09-146-053-5	Sequence 5, Appli
12	18	1.4	87350	3 US-08-781-891-79	Sequence 79, Appl
13	18	1.4	87350	4 US-09-618-166-79	Sequence 79, Appl
14	18	1.4	87543	4 US-09-791-211-3	Sequence 3, Appli
15	17	1.3	139	1 US-08-392-678-22	Sequence 22, Appl
16	17	1.3	139	1 US-08-457-304A-22	Sequence 22, Appl
17	17	1.3	139	1 US-08-456-701A-22	Sequence 22, Appl
18	17	1.3	139	1 US-08-684-932A-22	Sequence 22, Appl
19	17	1.3	307	4 US-09-313-294A-7088	Sequence 7088, Ap
20	17	1.3	436	4 US-09-397-787-288	Sequence 288, App
21	17	1.3	440	4 US-09-397-787-311	Sequence 311, App
22	17	1.3	571	4 US-09-620-312D-920	Sequence 920, App
23	17	1.3	597	2 US-08-332-766A-19	Sequence 19, Appl
24	17	1.3	864	4 US-09-252-991A-12026	Sequence 12026, A
25	17	1.3	870	4 US-09-252-991A-11957	Sequence 11957, A
26	17	1.3	924	4 US-09-252-991A-12055	Sequence 12055, A
27	17	1.3	1137	4 US-09-569-038A-104	Sequence 104, App

28 17 1.3 1177 2 US-08-841-349-11 Sequence 11, Appli  
29 17 1.3 1342 1 US-08-450-360-3 Sequence 3, Appli  
30 17 1.3 1345 3 US-08-975-762-1 Sequence 1, Appli  
31 17 1.3 1345 3 US-08-821-324-1 Sequence 1, Appli  
32 17 1.3 1345 3 US-09-295-028-1 Sequence 1, Appli  
33 17 1.3 1345 4 US-09-106-582-1 Sequence 1, Appli  
34 17 1.3 1399 4 US-09-996-243-27 Sequence 27, Appli  
35 17 1.3 1461 1 US-08-450-360-1 Sequence 1, Appli  
36 17 1.3 1700 2 US-08-428-131-1 Sequence 1, Appli  
37 17 1.3 1700 2 US-08-602-846-1 Sequence 1, Appli  
38 17 1.3 1700 3 US-09-078-596-1 Sequence 1, Appli  
39 17 1.3 2050 4 US-09-620-312D-761 Sequence 761, App  
40 17 1.3 3073 4 US-09-620-312D-279 Sequence 279, App  
41 17 1.3 3747 4 US-10-042-810-1 Sequence 1, Appli  
42 17 1.3 3837 4 US-10-042-810-3 Sequence 3, Appli  
43 17 1.3 4068 4 US-10-042-810-5 Sequence 5, Appli  
44 17 1.3 4772 4 US-09-484-970B-133 Sequence 133, App  
45 17 1.3 5889 4 US-09-402-929-3 Sequence 3, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103.840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "a" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 1.6%; Score 20; DB 3; Length 4403765;  
Best Local Similarity 100.0%; Pred. No. 3.4; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0

QY 500 AGCTGCGTGCATCTCTGCG 519  
|||||  
Db 7383 AGCTGCGTGCATCTCTGCG 7364

##### RESULT 2

US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103.840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match      1.6%; Score 20; DB 3; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      500 AGCTGGCTGCACCTCTCTGC 519
Db      7383 AGCTGGCTGCACCTCTCTGC 7364

RESULT 3
US-08-381-433A-3
; Sequence 3, Application US/08381433A
; Patent No. 5786203
; GENERAL INFORMATION:
; APPLICANT: Lovenberg, Timothy W.
; APPLICANT: Oltersdorf, Tilman
; APPLICANT: Liaw, Chen
; APPLICANT: Grigoriadis, Dimitri E.
; APPLICANT: DeSouza, Errol B.
; TITLE OF INVENTION: CORTICOTROPIN RELEASING FACTOR 2
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,433A
; FILING DATE: 31-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 690068.401C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANDBERRY
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 216..1449
US-08-381-433A-3

Query Match      1.5%; Score 19; DB 1; Length 1626;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      72 CCGGGGTGGCGGGGAGGC 90
Db      95 CCGGGGTGGCGGGGAGGC 113

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match      1.6%; Score 20; DB 3; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      500 AGCTGGCTGCACCTCTCTGC 519
Db      7383 AGCTGGCTGCACCTCTCTGC 7364

RESULT 4
US-09-157-177-128/c
; Sequence 128, Application US/09157177
; Patent No. 6090558
; GENERAL INFORMATION:
; APPLICANT: Butler, John M.
; APPLICANT: Li, Jia
; APPLICANT: Monforte, Joseph A.
; APPLICANT: Becker, Christopher H.
; TITLE OF INVENTION: DNA TYPING BY MASS SPECTROMETRY WITH POLYMORPHIC DNA
; TITLE OF INVENTION: REPEAT MARKERS
; FILE REFERENCE: GSTR:017/GSTR0179
; CURRENT APPLICATION NUMBER: US/09/157,177
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 60/059,415
; EARLIER FILING DATE: 1997-09-19
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 128
; LENGTH: 344
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-157-177-128

Query Match      1.4%; Score 18; DB 3; Length 344;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      982 TCTGGAACAAGGAGAAC 999
Db      42 TCTGGAACAAGGAGAAC 25

RESULT 5
US-09-280-116-14/c
; Sequence 14, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: thimet oligopeptidase
US-09-280-116-14

Query Match      1.4%; Score 18; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      522 GAAGAGAGAGGAGGTGGA 539
Db      302 GAAGAGAGAGGAGGTGGA 285

RESULT 6
US-09-535-008-44
; Sequence 44, Application US/09535008
; Patent No. 6465629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
```

FILE REFERENCE: 2318-259  
CURRENT APPLICATION NUMBER: US/09/535,008  
EARLIER FILING DATE: 2000-03-23  
EARLIER APPLICATION NUMBER: U.S. 60/125,806  
EARLIER FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 44  
LENGTH: 713  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-535-008-44

Query Match 1.4%; Score 18; DB 4; Length 713;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CCACCGCGCTCCAAAAGC 210  
|||||  
DB 256 CCACCGCGCTCCAAAAGC 273  
|||||

RESULT 7  
US-09-252-991A-9940/c  
Sequence 9940, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 9940  
LENGTH: 933  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-9940

Query Match 1.4%; Score 18; DB 4; Length 933;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 CCTTGACCGAGGAATCT 340  
|||||  
DB 451 CCTTGACCGAGGAATCT 434  
|||||

RESULT 8  
US-09-252-991A-10137  
Sequence 10137, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 10137  
LENGTH: 1023  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-10137  
Query Match 1.4%; Score 18; DB 4; Length 1023;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 CCTTGACCGAGGAATCT 340  
|||||  
DB 537 CCTTGACCGAGGAATCT 554  
|||||

RESULT 9  
US-09-023-942A-27/c  
Sequence 27, Application US/09023942A  
Patent No. 6479274  
GENERAL INFORMATION:  
APPLICANT: (US only) ANTALIS Toni Marie and HOOPER John David  
TITLE OF INVENTION: NOVEL MOLECULES  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 GARDEN CITY PLAZA  
CITY: GARDEN CITY  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,942A  
FILING DATE: 13-FEB-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P05101/97  
FILING DATE: 13-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P04022/97  
FILING DATE: 18-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: International PCT Application  
FILING DATE: 13-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: DIGILIO, FRANK S  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 11168  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742 4343  
TELEFAX: (516) 742 4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3866 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-023-942A-27

Query Match 1.4%; Score 18; DB 4; Length 3866;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 GTCGCCCCCTCGGGGCGAG 278  
|||||  
DB 465 GTCGCCCCCTCGGGGCGAG 448  
|||||

RESULT 10  
US-08-178-257-1  
Sequence 1, Application US/08178257

; Patent No. 6515205  
; GENERAL INFORMATION:  
; APPLICANT: LIEBERGSELL, MATTHIAS  
; APPLICANT: STEINBUCH, ALEXANDER  
; TITLE OF INVENTION: PRODUCTION OF POLYALKANONATE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3518  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/178,257  
; FILING DATE: 11-JAN-1994  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB92/01291  
; FILING DATE: 15-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9115245.4  
; FILING DATE: 16-JUL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOKULIS, PAUL N  
; REGISTRATION NUMBER: 16773  
; REFERENCE/DOCKET NUMBER: 9320/206058  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202/861-3000  
; TELEFAX: 202/822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4506 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Chromatium vinosum  
; US-08-178-257-1

Query Match 1.4%; Score 18; DB 4; Length 4506;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 GTAGCTTGGAGCGGCC 364  
DB 3222 GTAGCTTGGAGCGGCC 3239

RESULT 11  
US-09-146-053-5/c  
; Sequence 5, Application US/09146053A  
; Patent No. 6399349  
; GENERAL INFORMATION:  
; APPLICANT: Ryan, James W.  
; APPLICANT: Sprinkle, Terry Joe Curtis  
; APPLICANT: Venema, Richard C.  
; TITLE OF INVENTION: Human Aminopeptidase P Gene  
; FILE REFERENCE: MCG103  
; CURRENT APPLICATION NUMBER: US/09/146,053A  
; CURRENT FILING DATE: 1998-09-02  
; EARLIER APPLICATION NUMBER: 60/057,854  
; EARLIER FILING DATE: 1997-09-02  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 5  
; LENGTH: 44453  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-146-053-5

Query Match 1.4%; Score 18; DB 4; Length 44453;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 CCAGGCGCTCCCGGCA 230  
DB 11808 CCAGGCGCTCCCGGCA 11791

RESULT 12  
US-08-781-891-79  
; Sequence 79, Application US/08781891  
; Patent No. 6090620  
; GENERAL INFORMATION:  
; APPLICANT: Fu, Ying-Hui  
; APPLICANT: Yu, Chang-En  
; APPLICANT: Oshima, Junko  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Schellenberg, Gerald D.  
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
; TITLE OF INVENTION: WERNER'S SYNDROME  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781,891  
; FILING DATE: 27-DEC-1996  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6090620tenburg Ph.D., Carol  
; REGISTRATION NUMBER: 39,517  
; REFERENCE/DOCKET NUMBER: 240052.419  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 79:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 87350 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-781-891-79

Query Match 1.4%; Score 18; DB 3; Length 87350;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 TTCTGTTTCATTCAT 1166  
DB 34669 TTCTGTTTCATTCAT 34686

RESULT 13  
US-09-618-166-79  
; Sequence 79, Application US/09618166  
; Patent No. 6583112  
; GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui  
Yu, Chang-Su  
Oshima, Junko  
Mulligan, John T.  
Schellenberg, Gerald D.  
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
WERNER'S SYNDROME  
NUMBER OF SEQUENCES: 209  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed Intellectual Property Law Group  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/618,166  
FILING DATE: 17-Jul-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 240052.419C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87350 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 79:  
US-09-618-166-79

Query Match 1.4%; Score 18; DB 4; Length 87350;  
Best Local Similarity 100.0%; Pred.No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 TTCGTCATTCATTCAT 1166  
|||||  
DB 34669 TTCGTCATTCATTCAT 34686

RESULT 14  
US-09-791-211-3  
Sequence 3, Application US/09791211  
Patent No. 6448080  
GENERAL INFORMATION:  
APPLICANT: Donna T. Ward  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION  
FILE REFERENCE: RFS-0205  
CURRENT APPLICATION NUMBER: US/09/791,211  
CURRENT FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 90  
SEQ ID NO 3  
LENGTH: 87543  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 7421  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 7427  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 47291

LOCATION: 11609  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 12605  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 12742  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29370  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29422  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29979  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29980  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29981  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 30136  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 30140  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 31205  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 31206  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 31592  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 33095  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 33160  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 34066  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 34072  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 36816  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 39020  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 42164  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 42459  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46808  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46823  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46826  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 47291

OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 52786  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 52787  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 53384  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 54684  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59215  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59235  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59242  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 63290  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 66614  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68660  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68697  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68718  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68733  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68739  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 69785  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 79134  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 79198  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 86336  
OTHER INFORMATION: unknown  
OTHER INFORMATION:  
US-09-791-211-3

Query Match 1.4%; Score 18; DB 4; Length 87543;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 TTCTGTTTCATTTCAT 1166  
|||||  
Db 34862 TTCTGTTTCATTTCAT 34879

RESULT 15  
US-08-392-678-22  
Sequence 22, Application US/08392678  
Patent No. 5552281  
GENERAL INFORMATION:  
APPLICANT: Stashenko, Philip

APPLICANT: Li, Yi-Ping  
APPLICANT: Wuchertfennig, Anne L  
TITLE OF INVENTION: HUMAN OSTEOCLAST-SPECIFIC AND  
TITLE OF INVENTION: -RELATED GENES  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,678  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/045,270  
FILING DATE: 06 APR 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: FDC92-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-392-678-22

Query Match 1.3%; Score 17; DB 1; Length 139;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 TCTGCTTCCTGTGTGT 453  
|||||  
Db 94 TCTGCTTCCTGTGTGT 110

Search completed: December 6, 2003, 01:46:43  
Job time : 117 secs

